



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 99157

TO: Ulrike Winkler
Location: CM1/8D09&8E12
Art Unit: 1648
Friday, July 18, 2003

Case Serial Number: 10/054968

From: Edward Hart
Location: Biotech-Chem Library
CM1-6B02
Phone: 305-9203

edward.hart@uspto.gov

Search Notes

Examiner Winkler,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

antibody → ABC transporter
immunogenic fragment seq 6, 7 and 8.
→ Need Ab to recognize seq 10 6, 7 and 8.

Priority date
July 6, 1996

STIC-Biotech/ChemLib

99157

From: Winkler, Ulrike
Sent: Friday, July 18, 2003 8:49 AM
T : STIC-Biotech/ChemLib

STIC

Please search SEQ ID 6, 7, 8 of 10/054968.

Thanks, Ulrike

Ulrike Winkler, Ph.D.
Patent Examiner
Crystal Mall 1, 8D09/ Mail Box 8E12
1911 South Clark Place
Arlington, VA 22202
tel. 703-308-8294
fax. 703-308-4426

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 7/18/03
Searcher Prep/Review: 7/18/03
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 3
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 02
WWW/Internet: _____
Other (specify): _____



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or *contact:*

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 – Circ. Desk



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:59:34 ; Search time 15.1579 Seconds
(without alignments)
50.738 Million cell updates/sec

Title: US-10-054-968-8

Perfect score: 35

Sequence: 1 RVALAGVL 8

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	183	2 T44413	ABC transporter (A
2	35	100.0	276	2 E69742	ABC transporter (A
3	35	100.0	279	2 AH1775	ABC transporter (A
4	35	100.0	279	2 A11399	ABC transporter (A
5	35	100.0	281	2 G97281	ABC-type transport
6	35	100.0	282	2 E83670	ABC transporter (A
7	35	100.0	288	2 AG1775	ABC transporter (A
8	35	100.0	288	2 AH1399	ABC transporter (A
9	34	97.1	277	2 T45265	ABC transporter (A
10	34	97.1	278	2 T44412	ABC transporter (A
11	34	97.1	279	2 G64435	cobalt transporter (A
12	34	97.1	279	2 C95259	ABC transporter (A
13	34	97.1	279	2 F98124	hypothetical prote
14	34	97.1	288	2 F86659	ABC transporter AT
15	34	97.1	303	2 S62814	histidine transpo
16	34	97.1	304	2 I64219	membrane transpo
17	34	97.1	311	2 A69095	ABC transporter A
18	34	97.1	433	2 D82879	ABC transporter UU
19	33	94.3	150	2 D84173	archaeal histone A
20	33	94.3	281	2 E69751	ABC transporter (A
21	32	91.4	275	2 D95259	ABC transporter (A
22	32	91.4	275	2 G98124	hypothetical prote
23	32	91.4	280	2 G69043	hypothetical prote
24	32	91.4	286	2 F97281	ABC-type transport
25	32	91.4	304	2 C90560	ABC transporter at
26	31	88.6	180	2 A42283	hypothetical prote
27	31	88.6	268	2 AG1225	cobalt transport A
28	31	88.6	268	2 A11578	cobalt transport A
29	31	88.6	271	2 AG0757	probable cobalt tr

30	31	88.6	277	2 E86659	ABC transporter AT
31	31	88.6	287	2 E69778	conserved hypothet
32	31	88.6	344	2 E75441	peptidyl-prolyl ci
33	31	88.6	364	2 G70451	conserved hypothet
34	31	88.6	586	2 D69250	RNase L inhibitor
35	31	88.6	630	2 C90470	hypothetical prote
36	30	85.7	62	2 E84200	hypothetical prote
37	30	85.7	178	2 S77760	probable ABC-type
38	30	85.7	242	2 F84315	cobalt transport A
39	30	85.7	270	2 T05997	hypothetical prote
40	30	85.7	328	2 H81229	DNA-directed RNA p
41	30	85.7	360	2 G72673	hypothetical prote
42	30	85.7	473	2 AC2251	hypothetical prote
43	30	85.7	560	2 A95056	ABC transporter. A
44	30	85.7	560	2 F97925	hypothetical prote
45	30	85.7	565	2 E86665	ABC transporter AT

ALIGNMENTS

RESULT 1
T44413
ABC transporter (ATP-binding protein) homolog ybaE [imported] - Bacillus halodurans
C:Species: Bacillus halodurans
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 21-Jul-2000
C:Accession: T44413
R:Takami, H.; Takaki, Y.; Nakasone, K.; Hirama, C.; Inoue, A.; Horikoshi, K.
Biosci. Biotechnol. Biochem. 63, 452-455, 1999
A:Title: Sequence analysis of a 32-kb region including the major ribosomal protein ge
A:Reference number: Z22756; MUID:99209008; PMID:10192928
A:Accession: T44413
A:Status: preliminary; translated from GB/EML/DDBU
A:Molecule type: DNA
A:Residues: 1-183 <TAK>
A:Cross-references: EMBL:AB017508; NID:94512395; PIDM:BAAT5301.1; PID:94512434
A:Experimental source: strain C-125
C:Genetics:
A:Note: ybaE

Query Match 100.0%; Score 35; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVALAGVL 8
DB 42 RVALAGVL 49

RESULT 2
E69742
ABC transporter (ATP-binding protein) homolog ybaE - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Feb-2001
C:Accession: E69742
R:Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Acevedo, V.; Ber
C.; Brox, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Enlian, K.D.; Errington, J.; Fabbri, C.; Ferrari,
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fuma, S.; Gallizi, A.; Gal
lechi, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Larding
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scan
A:Authors: Schleich, S.; Schroeter, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; S
akenuchi, M.; Tamakoshi, A.; Tanaka, T.; Terstra, P.; Tognoni, A.; Tosato, V.; Uchi
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshi
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subt
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: E69742
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-276 <KUN>
 A:Cross-references: GB:299104; GB:AL009126; NID:g2632267; PIDN:CAB11922.1; PID:el182079;
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: ybbE
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 C:Keywords: ATP; nucleotide binding; P-loop
 F:10-209/Domain: ATP-binding cassette homology <ABC>
 F:27-34/Region: nucleotide-binding motif A (P-loop)

Query Match 100.0%; Score 35; DB 2; Length 276;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 RVALAGVL 8
 |||||
 Db 140 RVALAGVL 147

RESULT 3
 ABC transporter (ATP-binding protein) homolog lin2750 [imported] - Listeria innocua (str
 C:Species: Listeria innocua
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
 C:Accession: AH1775
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 : Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlan, K.D.; Fsihl, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maltournam, A.; Ma
 ok, C.; Schluter, T.; Simoes, N.; Tlertez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
 A:Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AH1775
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-279 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC97976.1; PID:g16415286; GSPDB:GN00178
 A:Experimental source: strain C14p11262
 C:Genetics:
 A:Gene: lin2750
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 100.0%; Score 35; DB 2; Length 279;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 RVALAGVL 8
 |||||
 Db 148 RVALAGVL 155

RESULT 4
 A11399
 ABC transporter (ATP-binding protein) homolog lmo2601 [imported] - Listeria monocytogenes
 C:Species: Listeria monocytogenes
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
 C:Accession: A11399
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 : Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlan, K.D.; Fsihl, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maltournam, A.; Ma
 ok, C.; Schluter, T.; Simoes, N.; Tlertez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
 A:Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: A11399
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-279 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAD00679.1; PID:g16412089; GSPDB:GN00177
 A:Experimental source: strain EGD-e

C:Genetics:
 A:Gene: lmo2601
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 100.0%; Score 35; DB 2; Length 279;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 RVALAGVL 8
 |||||
 Db 148 RVALAGVL 155

RESULT 5
 G97281
 ABC-type transporter, ATPase component (cobalt transporters subfamily) CAC3102 [import
 C:Species: Clostridium acetobutylicum
 C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C:Accession: G97281
 R:Mollng, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
 : Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: G97281
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-281 <KUR>
 A:Cross-references: GB:AE001437; PIDN:AAK81042.1; PID:g15026167; GSPDB:GN00168
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC3102

Query Match 100.0%; Score 35; DB 2; Length 281;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 RVALAGVL 8
 |||||
 Db 153 RVALAGVL 160

RESULT 6
 E83670
 ABC transporter (ATP-binding protein) BH0165 [imported] - Bacillus halodurans (strain
 C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: E83670
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: E83670
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-282 <STO>
 A:Cross-references: GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BA03884.1; GSPDB:G
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH0165
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 100.0%; Score 35; DB 2; Length 282;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 RVALAGVL 8
 |||||
 Db 141 RVALAGVL 148

RESULT 7
 AG1775

ABC transporter (ATP-binding protein) homolog lin2749 [imported] - *Listeria innocua* (str
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AG1775
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Me
Ok, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehländ,
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1775
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-288 <GLA>
A:Cross-References: GB:AL592022; PIDN:CA97975.1; PID:g16415285; GSPDB:GN00178
A:Experimental source: strain C11p11262
C:Genetics:
A:Gene: lin2749
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Query Match 100.0%; Score 35; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
DB 153 RVAIAGVL 160

RESULT 8
AH1399
ABC transporter (ATP-binding protein) homolog lmo2600 [imported] - *Listeria monocytogenes*
C:Species: *Listeria monocytogenes*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AH1399
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
Ok, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehländ,
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1399
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-288 <GLA>
A:Cross-References: GB:NC_003210; PIDN:CAD00678.1; PID:g16412088; GSPDB:GN00177
A:Experimental source: strain EED-e
C:Genetics:
A:Gene: lmo2600
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Query Match 100.0%; Score 35; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
DB 153 RVAIAGVL 160

RESULT 9
T45265
cobalt transport ATP-binding protein O homolog [imported] - *Methanobacterium thermoautot*
C:Species: *Methanobacterium thermoautotrophicum*
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T45265
R:Eberhardt, S.; Korn, S.; Lottspeich, F.; Becher, A.
J. Bacteriol. 179, 2938-2943, 1997
A:Title: Biosynthesis of riboflavin: an unusual riboflavin synthase of *Methanobacterium*

A:Reference number: 222951; MUID:97284497; PMID:9139911
A:Accession: T45265
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-277 <EBD>
A:Cross-References: EMBL:X94292; NID:g1419074; PIDN:CAA63958.1; PID:g1419078
A:Experimental source: strain Marburg
C:Genetics:
A:Gene: cbio
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Query Match 97.1%; Score 34; DB 2; Length 277;
Best Local Similarity 87.5%; Pred. No. 9.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
DB 146 RVAIAGVL 153

RESULT 10
T44412
ABC transporter (ATP-binding protein) BH0164 [imported] - *Bacillus halodurans* (strain
C:Species: *Bacillus halodurans*
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 15-Jun-2001
C:Accession: T44412; D83670
R:Takami, H.; Takaki, Y.; Nakasone, K.; Hirama, C.; Inoue, A.; Horikoshi, K.
Biosci. Biotechnol. Biochem. 63, 452-455, 1999
A:Title: Sequence analysis of a 32-kb region including the major ribosomal protein ge
A:Reference number: 222756; MUID:99209008; PMID:10192928
A:Accession: T44412
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-278 <TAK>
A:Cross-References: EMBL:AB017508; NID:94512395; PIDN:BAV75300.1; PID:94512433
A:Experimental source: strain C-125
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a
A:Reference number: AB650; MUID:20512582; PMID:11058132
A:Accession: D83670
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-278 <STO>
A:Cross-References: GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BA03883.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0164
A:Note: ybxa
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Query Match 97.1%; Score 34; DB 2; Length 278;
Best Local Similarity 87.5%; Pred. No. 9.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
DB 148 RVAIAGVL 155

RESULT 11
G64435
cobalt transport ATP-binding protein O homolog - *Methanococcus jannaschii*.
C:Species: *Methanococcus jannaschii*
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Feb-2001
C:Accession: G64435
R:Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glöck,
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Bordovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannasc*
A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: G64435
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-279 <BUL>
 A:Cross-references: GB:U67551; GB:U77117; NID:g1591728; PIDN:AAB99089.1; PID:g1591732; T
 C:Genetics:
 A:Map position: REV1027976-1027137
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 C:Keywords: ATP; nucleotide binding; P-loop
 F:20-215/Domain: ATP-binding cassette homology <ABC>
 F:37-44/Region: nucleotide-binding motif A (P-loop)

Query Match 97.1%; Score 34; DB 2; Length 279;
 Best Local Similarity 87.5%; Pred. No. 9.9;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVALAGVL 8
 |||||:1
 DB 147 RVALAGIL 154

RESULT 12

ABC transporter, ATP-binding protein SP2220 [imported] - Streptococcus pneumoniae (strai
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 17-May-2002
 C:Accession: C95259
 R:Telletlin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: C95259
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-279 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK76268.1; PID:g14973730; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP2220
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 97.1%; Score 34; DB 2; Length 279;
 Best Local Similarity 87.5%; Pred. No. 9.9;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVALAGVL 8
 |||||:1
 DB 153 RVALAGIL 160

RESULT 13

hypothetical protein ABC-NBP [imported] - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 17-May-2002
 C:Accession: F98124
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
 y, R.; Leblang, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: F98124
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-279 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAL00827.1; PID:g15459731; GSPDB:GN00174
 C:Genetics:
 A:Gene: ABC-NBP

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 Query Match 97.1%; Score 34; DB 2; Length 279;
 Best Local Similarity 87.5%; Pred. No. 9.9;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVALAGVL 8
 |||||:1
 DB 153 RVALAGIL 160

RESULT 14

F86659
 ABC transporter ATP-binding protein yche [imported] - Lactococcus lactis subsp. lacti
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: F86659
 R:Boletín, A.; Winkler, P.; Mager, S.; Jallón, O.; Malame, K.; Weissenbach, J.; En
 genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
 A:Reference number: A86625; MUID:21235186; PMID:11337471
 A:Accession: F86659
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-288 <STO>
 A:Cross-references: GB:AE005176; PID:g12723140; PIDN:AAK04376.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: yche
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 97.1%; Score 34; DB 2; Length 288;
 Best Local Similarity 87.5%; Pred. No. 10;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVALAGVL 8
 |||||:1
 DB 152 RVALAGIL 159

RESULT 15

S62814
 histidine transport ATP-binding protein hisP - Mycoplasma pneumoniae (strain ATCC 293
 N:Alternate names: glutamine transport protein GlnG homology; hypothetical protein G79.
 C:Species: Mycoplasma pneumoniae
 A:Variety: ATCC 29342
 C:Date: 23-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 02-Feb-2001
 C:Accession: S62814; S73963
 R:Hilbert, H.; Himmelreich, R.; Plagens, H.; Herrmann, R.
 Nucleic Acids Res. 24, 628-639, 1996
 A:Title: Sequence analysis of 56 kb from the genome of the bacterium Mycoplasma pneum
 A:Reference number: S62797; MUID:96177562; PMID:8604303
 A:Accession: S62814
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-303 <HTL>

A:Cross-references: EMBL:U34795; NID:g1215663; PIDN:AAC34687.1; PID:g1215694
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
 R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996
 A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumon
 A:Reference number: S73327; MUID:97105885; PMID:8948633
 A:Accession: S73963
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-303 <HTL>

A:Cross-references: EMBL:AE000061; GB:U00089; NID:g1674336; PIDN:AAB96285.1; PID:g167
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C:Genetics:
 A:Gene: hisP
 A:Genetic code: SGC3
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 C:Keywords: ATP; nucleotide binding; P-loop

F:37-236/Domain: ATP-binding cassette homology <ABC>
F:54-61/Region: nucleotide-binding motif A (P-loop)

Query Match 97.1%; Score 34; DB 2; Length 303;
Best Local Similarity 87.5%; Pred. No. 11;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RVAIAGVL 8
|||||:

DB 168 RVAIAGIL 175

Search completed: July 18, 2003, 16:04:28
Job time : 17.1579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:57:45 ; Search time 8.8421 Seconds
(without alignments)
37.526 Million cell updates/sec

Title: US-10-054-968-8
Perfect score: 35
Sequence: 1 RVALAGVL 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	97.1	303	Y180_MYCPN	Q50293 mycoplasma
2	34	97.1	304	Y180_MYCPN	P47426 mycoplasma
3	33	94.3	281	YBXA_BACSU	P40735 bacillus su
4	31	88.6	271	CHIO_SALTY	O05596 salmonella
5	30	85.7	258	DAB2_RHITO	P58211 rhizobium l
6	30	85.7	328	KFOA_NEIMA	O9J106 neisseria m
7	29	82.9	274	Y179_MYCPN	P47425 mycoplasma
8	29	82.9	327	Y179_MYCPN	O50294 mycoplasma
9	29	82.9	327	PXA2_SALTY	P58716 salmonella
10	29	82.9	327	PXA2_SALTY	P58718 salmonella
11	29	82.9	485	PMDA_ECOLI	P24241 escherichia
12	29	82.9	531	NOON_MTCU	O53308 mycobacteri
13	28	80.0	161	LSPA_SYNY3	P73540 synchocyst
14	28	80.0	233	PYRE_COLGR	P35788 colletoctric
15	28	80.0	265	FLIR_TREPA	P74932 treponema p
16	28	80.0	321	PXA2_RHIME	O92x16 rhizobium m
17	28	80.0	355	CAHC_RABYT	O9mz30 oryctolagus
18	28	80.0	469	IARL_ARATH	O9m647 arabidopsis
19	28	80.0	557	MAOX_PIG	O29558 sus scrofa
20	28	80.0	604	MAON_HUMAN	O16798 homo sapien
21	28	80.0	808	SYEB_BACHD	O9K896 bacillus ha
22	28	80.0	810	SYEB_SYNY3	P74296 synchocyst
23	28	80.0	956	GLK4_HUMAN	O16099 homo sapien
24	28	80.0	956	GLK4_RAT	O01812 rattus norv
25	27	77.1	95	CILG_HAEIN	P44446 haemophilus
26	27	77.1	154	YV16_ARCEU	O29246 archaeoglob
27	27	77.1	169	GSPH_XANCP	P31736 xanthomonas
28	27	77.1	185	DABP_KLEPN	P45415 klebsiella
29	27	77.1	273	DABP_ECO57	P58209 escherichia
30	27	77.1	273	DABP_ECOLI	O84036 escherichia
31	27	77.1	273	DABP_SALTY	O82rx8 salmonella
32	27	77.1	273	DABP_SALTY	O82rx8 salmonella
33	27	77.1	296	YWFH_BACSU	P39649 bacillus su

34	27	77.1	362	1	OPSR_CHICK	P22329 gallus gall
35	27	77.1	362	1	PDXA_ANASP	P58711 anabaena sp
36	27	77.1	367	1	Y046_MYCTU	P71703 mycobacteri
37	27	77.1	369	1	Y046_MYCTU	O57240 mycobacteri
38	27	77.1	372	1	COQ2_YEAST	P32378 saccharomyc
39	27	77.1	388	1	ALR3_RHITO	O981h7 rhizobium l
40	27	77.1	474	1	CRPI_SYNP7	P26794 synchococc
41	27	77.1	497	1	HXK1_TOBAC	O9sek2 nicotiana t
42	27	77.1	528	1	NIFK_FRAL	O57118 frankia aln
43	27	77.1	549	1	YJCE_ECOLI	P32703 escherichia
44	27	77.1	615	1	MUTA_MYCTU	P71773 mycobacteri
45	27	77.1	623	1	MAOM_AMAHP	P37224 amarantus

ALIGNMENTS

RESULT 1
ID Y180_MYCPN STANDARD: PRT: 303 AA.
AC O50293:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical ABC transporter ATP-binding protein MG180 homolog (G9_orf303).
DE
GN MPN194 OR MP637.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE=96177562; PubMed=8604303;
RA Hilbert H., Himmelsreith R., Plagens H., Herrmann R.;
RT "Sequence analysis of 56 kb from the genome of the bacterium Mycoplasma pneumoniae comprising the dnaA region, the atp operon and a cluster of ribosomal protein genes.";
RT Nucleic Acids Res. 24:628-639(1996).
RL [2]
RN
RP
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelsreith R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.";
RT Nucleic Acids Res. 24:4420-4449(1996).
RL
CC -I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC
CC
CC EMBL: U34795; AAC3687.1; -;
CC EMBL: AE000061; AAB96285.1; -;
CC HSSP: 058663; 1G6H.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; AAA_TRANSPOR.
DR Pfam: PF00005; ABC_tran.1.
DR ProDom: P000006; ABC_transport.1.
DR SMART: SM00382; AAA.1.
DR PROSITE: PS00211; ABC_TRANSPORTER.1.
KW Hypothetical protein: ATP-binding; Transport; Complete proteome.
FT NP_BIND 54 61 ATP (POTENTIAL).
SQ SEQUENCE 303 AA; 34357 MW; CQEDD5F81BFA19 CRC64;
Query Match 97.1%; Score 34; DB 1; Length 303;
Best Local Similarity 87.5%; Pred. No. 6.6;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
168 RVAIAGIL 175

RESULT 2

ID Y180_MYCGE STANDARD; PRT; 304 AA.
AC P47426; Q49207;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical ABC transporter ATP-binding protein MG180.
GN MG180.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RA MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann J.D., Smith R.L., Peterson S.N., White O.,
RA Fritchman J.E., Weidman J.F., Small K.V., Sandusky M.,
RA Nguyen D.T., Ullrich T.R., Saudek D.M., Phillips C.A.,
RA Tomb J.F., Dougherty B.A., Burt K.F., Hu P.-C.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RA "The minimal gene complement of Mycoplasma genitalium";
RL Science 270:397-403(1995).
[2]
RP SEQUENCE OF 10-83 AND 262-304 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RA MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Burt K.F., Hutchison C.A. III;
RA "A survey of the Mycoplasma genitalium genome by using random
RT sequencing";
RL J. Bacteriol. 175:7918-7930(1993).
[3]
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U39697; AAC71399.1; -;
CC EMBL: U01754; AAD10568.1; -;
CC EMBL: U01750; AAD10564.1; -;
CC TIGR: MG180; -;
CC InterPro: IPR003593; AAA_ATPase.
CC InterPro: IPR003439; ABC_transporter.
CC Pfam: PF00005; ABC_tran. 1.
CC ProDom: PD000006; ABC_transporter. 1.
CC SMART: SM00382; AAA. 1.
CC PROSITE: PS00211; ABC_TRANSPORTER. 1.
CC Hypothetical protein; ATP-binding; Transport; Complete proteome.
FT NP_BIND 54 61 ATP (POTENTIAL).
FT CONFLICT 11 14 LKAD -> FKSS (IN REF. 2).
SQ SEQUENCE 304 AA; 34385 MW; 32B05A81138DCAB2 CRC64;

Query Match 97.1%; Score 34; DB 1; Length 304;
Best Local Similarity 87.5%; Pred. No. 6.6;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
168 RVAIAGIL 175

RESULT 3

ID YBXA_BACSU STANDARD; PRT; 281 AA.
AC P40735;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical ABC transporter ATP-binding protein ybxa.
GN YBXA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA MEDLINE=97124188; PubMed=8969501;
RA Yasunoto K., Liu H., Jeong S.M., Ohashi Y., Kakinuma S., Tanaka K.,
RA Kawamura F., Yoshikawa H., Takahashi H.;
RA "Sequence analysis of a 50 kb region between spoOH and trnH on the
RT Bacillus subtilis chromosome";
RL Microbiology 142:3039-3046(1996).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA MEDLINE=96044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Besterio M.G., Bessieres P., Bilotin A., Borchert S.,
RA Boursier L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Enlrich S.D., Emmerson P.T.,
RA Enlian K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallazzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holstapfel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,
RA Kodayashiki Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Meisel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nockack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudga B., Park S.H.,
RA Parro V., Pohl T.M., Portetle D., Portwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche B., Rose M., Sadleir Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Seliguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Solido B.,
RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
RA Takenuchi M., Tanakoshi A., Tanaka T., Terpestra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambolt R., Wedler E., Wedler H., Weltenegeger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RA "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256(1997).
[3]
RP SEQUENCE OF 1-31 FROM N.A.
RC MEDLINE=89213940; PubMed=2496109;
RA Boylan S.A., Suh J.-W., Thomas S.M., Price C.W.;
RA "Gene encoding the alpha core subunit of Bacillus subtilis RNA
RT polymerase is cotranscribed with the genes for Initiation factor 1
RT J. Bacteriol. 171:2553-2562(1989).
[4]
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```

CC -----
CC EMBL: D64126; BAA10983.1; -.
CC DR EMBL: 299104; CAB11921.1; -.
CC DR EMBL: M26414; AAA22219.1; -.
CC DR PIR: G32307; G32307.
CC DR Subtilisin; Bg11067; ybxa.
CC DR InterPro: IPR003593; AAA_Atpase.
CC DR InterPro: IPR003439; ABC_transporter.
CC DR Pfam: PF000005; ABC_tran.1.
CC DR Prodom: PD000006; ABC_transporter; 1.
CC DR SMART; SM00382; AAA; 1.
CC DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
CC DR Hypothetical protein; ATP-binding; Transport; Complete proteome.
CC NP_BIND 42 49 ATP (POTENTIAL).
CC FT SEQUENCE 281 AA; 31461 MW; 3CB4B8DB2340F68E CRC64;
CC SO
CC
CC Query Match 94.3%; Score 33; DB 1; Length 281;
CC Best Local Similarity 87.5%; Pred. NO. 9.9;
CC Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC 1 RVAIAGVL 8
CC 150 RVAIAGVI 157
CC
CC RESULT 4
CC CBIO_SALTY
CC ID CBIO_SALTY STANDARD; PRT; 271 AA.
CC AC 005596;
CC DT 01-JUN-1994 (Rel. 29, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Cobalt transport ATP-binding protein cbio.
CC GN CBIO OR STM2020.
CC OS Salmonella typhimurium.
CC OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC OC Salmonella.
CC OX NCBI_TaxID=602;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-LT2;
CC MEDLINE=93273696; PubMed=8501034;
CC RA Ruch J.R., Lawrence J.G., Rubenfield M., Kieffer-Higgins S.,
CC RA Church G.M.;
CC RT "Characterization of the cobalamin (vitamin B12) biosynthetic genes
CC RT of Salmonella typhimurium.";
CC RL J. Bacteriol. 175:3503-3516(1993).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-LT2 / SGSC1412 / ATCC 700720;
CC MEDLINE=21534948; PubMed=11677609;
CC RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
CC RA Courtney L., Portwolik S., All J., Dante M., Du F., Hou S., Layman D.,
CC RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
CC RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
CC RA Waterston R., Wilson R.K.;
CC RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
CC RT LT2.";
CC RL Nature 413:852-856(2001).
CC CC -1- FUNCTION: MAY BE INVOLVED WITH COBALT TRANSPORT IN ASSOCIATION
CC CC WITH COBALAMIN BIOSYNTHESIS.
CC CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: L12006; AAA27267.1; -.

```

```

CC -----
CC DR EMBL: AE008789; AAL20924.1; -.
CC DR StGene: SG10047; cbio.
CC DR InterPro: IPR003593; AAA_Atpase.
CC DR InterPro: IPR003439; ABC_transporter.
CC DR Pfam: PF000005; ABC_tran.1.
CC DR Prodom: PD000006; ABC_transporter; 1.
CC DR SMART; SM00382; AAA; 1.
CC DR TIGRfams; TIGR01166; cbio; 1.
CC DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
CC DR Cobalamin biosynthesis; Cobalt transporter; Transport; Inner membrane;
CC KW ATP-binding; Complete proteome.
CC FT NP_BIND 34 75 L -> R (IN REF. 1).
CC FT CONFLICT 75 75 L -> G (IN REF. 1).
CC FT CONFLICT 108 108 V -> G (IN REF. 1).
CC FT SEQUENCE 271 AA; 30147 MW; D5D7F707DA76CB6C CRC64;
CC SQ
CC
CC Query Match 88.6%; Score 31; DB 1; Length 271;
CC Best Local Similarity 87.5%; Pred. NO. 25;
CC Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC
CC 1 RVAIAGVL 8
CC 144 RVAIAGAL 151
CC
CC RESULT 5
CC DAB2_RHIL0
CC ID DAB2_RHIL0 STANDARD; PRT; 258 AA.
CC AC P58211;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Dihydropicolinate reductase 2 (EC 1.3.1.26) (DHPR 2).
CC GN DAB2 OR MUR7948.
CC GN Rhizobium loti (Mesorhizobium loti).
CC OS Rhizobium loti (Mesorhizobium loti).
CC OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
CC OC Phyllobacteriaceae; Mesorhizobium.
CC OX NCBI_TaxID=381;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-MAFF30309;
CC MEDLINE=21082930; PubMed=11214968;
CC RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
CC RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
CC RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
CC RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
CC RA Takeuchi C., Yamada M., Tabata S.;
CC RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
CC RT Mesorhizobium loti.";
CC RL DNA Res. 7:331-338(2000).
CC CC -1- CATALYTIC ACTIVITY: 2,3,4,5-tetrahydropicolinate + NAD(P)(+) =
CC CC 2,3-dihydropicolinate + NAD(P)H.
CC CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
CC CC semialdehyde; second step.
CC CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC CC -1- SIMILARITY: BELONGS TO THE DIHYDRODIPICOLINATE REDUCTASE FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: AP003013; BAB53612.1; ALT_INIT.
CC CC HSSP; P04036; IDRW.
CC DR InterPro: IPR000846; DapB.
CC DR Pfam: PF01113; DapB; 1.
CC DR Prodom: PD004105; DapB; 1.
CC DR TIGRfams; TIGR00036; dapB; 1.
CC DR PROSITE; PS01298; DAPB; 1.
CC KW Diaminopimelate biosynthesis; Lysine biosynthesis; Oxidoreductase;

```

KW NADP, Complete proteome.
 SQ SEQUENCE 258 AA; 26726 MW; 7D0F44EA0067A7F0 CRC64;
 Query Match 85.7%; Score 30; DB 1; Length 258;
 Best Local Similarity 75.0%; Pred. No. 39;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVAIAGVL 8
 Db 7 RIAIAGAL 14

RESULT 6
 RPOA_NEIMA STANDARD; PRT; 328 AA.
 ID RPOA_NEIMA
 AC Q9JRO6;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE DNA-directed RNA polymerase alpha chain (EC 2.7.7.6) (Transcriptase
 alpha chain) (RNA polymerase alpha subunit).
 GN RPOA OR NMA0103 OR NMB0168.
 OS Neisseria meningitidis (serogroup A), and
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699, 491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z2491 / Serogroup A; Serotype 4A;
 RX MEDLINE=20222555; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
 RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Symmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 meningitidis Z2491."
 RL Nature 404:502-506(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Nelson J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Eison W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Cline H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Yamathayan J.,
 RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappunli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58."
 RL Science 287:1809-1815(2000).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 (RNA)(N).
 CC -1- SUBUNIT: CONSISTS OF A SIGMA FACTOR AND THE RNAP CORE ENZYME WHICH
 IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, 1 BETA' CHAIN AND 1
 OMEGA CHAIN (BY SIMILARITY).
 CC -1- DOMAIN: THE AMINO-TERMINAL PORTION IS INVOLVED IN THE ASSEMBLY OF
 CORE RNAP, WHEREAS THE C-TERMINAL IS INVOLVED IN INTERACTION WITH
 TRANSCRIPTIONAL REGULATORS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE ALPHA CHAIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: AL162752; CAB83418.1;
 DR EMBL: AE002374; AAF40625.1;
 DR HSPB: P00574; ICCO.
 DR TIGR: NMB0168;
 DR InterPro: IPR001700; RNA_pol_A_bac.
 DR Pfam: PF01000; RNA_pol_A_bac; 1.
 DR Pfam: PF03118; RNA_pol_A_CTD; 1.
 DR ProDom: PD001179; RNA_pol_A_bac; 1.
 KW Transferrase; Transcription; DNA-directed RNA polymerase;
 KW Complete proteome.
 SQ SEQUENCE 328 AA; 36076 MW; 893A4C39571DD34D CRC64;
 Query Match 85.7%; Score 30; DB 1; Length 328;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAIAAGVL 8
 Db 59 VAIAAGVL 65

RESULT 7
 Y179_MYCGE STANDARD; PRT; 274 AA.
 ID Y179_MYCGE
 AC P47425;
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical ABC transporter ATP-binding protein MG179.
 GN MG179.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fitchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.F., Dougherty B.A., Bott K.F., Hu P.C., Luetter T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium."
 RL Science 270:397-403(1995).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL: U39697; AAC71398.1;
 DR TIGR: MG179;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transporter.
 DR Pfam: PF00005; ABC_tran; 1.
 DR ProDom: PD000006; ABC_transporter; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW Hypothetical protein; ATP-binding; Transport; Complete proteome.
 FT NP_BIND 41 48 ATP (POTENTIAL).
 SQ SEQUENCE 274 AA; 30544 MW; 5A70BD912AF054B1 CRC64;
 Query Match 82.9%; Score 29; DB 1; Length 274;
 Best Local Similarity 87.5%; Pred. No. 66;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
 |||||
 DB 149 RVAIASVL 156

RESULT 8

ID Y179_MYCPN STANDARD: PRT: 274 AA.
 AC 050294;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical ABC transporter ATP-binding protein MG179 homolog
 DE (G79_orf274).
 GN MPN193 OR MP638.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 29342 / M129;
 RX MEDLINE=96177562; PubMed=8604303;
 RA Hilbert H., Himmelfreich R., Plagens H., Herrmann R.;
 RT "Sequence analysis of 56 kb from the genome of the bacterium
 RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and
 RT a cluster of ribosomal protein genes.";
 RL Nucleic Acids Res. 24:628-639(1996).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelfreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae.";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1 SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U34795; AAC43688.1; -;
 DR EMBL: AE000061; AAB96286.1; -;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transportr.
 DR Pfam: PF00005; ABC_tran.1.
 DR ProDom: PD000006; ABC_transportr.1.
 DR SMART: SM00382; AAA.1.
 DR PROSITE: PS00211; ABC_TRANSPORTER.1.
 DR Hypothetical protein: ATP-binding; transport; Complete proteome.
 FT NP_BIND 42 49 ATP (POTENTIAL).
 SQ SEQUENCE 274 AA; 30535 MW; 83591C25C7B7552 CRC64;

Query Match 82.9%; Score 29; DB 1; Length 274;
 Best Local Similarity 87.5%; Pred. No. 66;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
 |||||
 DB 150 RVAIASVL 157

RESULT 9

ID PXA2_SALTI STANDARD: PRT: 327 AA.
 AC P58716;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 4-hydroxythreonine-4-phosphate dehydrogenase 2 (EC 1.1.1.262) (4-
 DE (phosphohydroxy)-L-threonine dehydrogenase 2).
 GN PDXA2 OR STY0185.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=601;
 RN [1]

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parthall J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Fellwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jajels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 CC -1 FUNCTION: Catalyzes the NAD-dependent oxidation of 4-
 CC (phosphohydroxy)-L-threonine (HTP) into 2-amino-3-oxo-4-
 CC (phosphohydroxy)butyric acid which spontaneously decarboxylate to
 CC form L-amino-3-(phosphohydroxy)propan-2-one (3-amino-2-oxopropyl
 CC phosphate) (by similarity).
 CC -1 CATALYTIC ACTIVITY: 4-(phosphonooxy)-threonine + NAD(+) = 2-
 CC amino-3-oxo-4-phosphonooxybutyrate + NADH.
 CC -1 PATHWAY: De novo synthesis of pyridoxine (vitamin B6) and
 CC pyridoxal phosphate.
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1 SIMILARITY: BELONGS TO THE PDXA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AL627265; CAD01321.1; -;
 DR InterPro: IPR00557; PDXA.1.
 DR TIGRFAMs: TIGR00557; PDXA.1.
 KW Pyridoxine biosynthesis; Oxidoreductase; NAD; Complete proteome.
 SQ SEQUENCE 327 AA; 35041 MW; F7F53DEB5326FEB CRC64;

Query Match 82.9%; Score 29; DB 1; Length 327;
 Best Local Similarity 71.4%; Pred. No. 78;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGV 7
 ||:|
 DB 204 RIAVAGV 210

RESULT 10

ID PXA2_SALTY STANDARD: PRT: 327 AA.
 AC P58718;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 4-hydroxythreonine-4-phosphate dehydrogenase 2 (EC 1.1.1.262) (4-
 DE (phosphohydroxy)-L-threonine dehydrogenase 2).
 GN PDXA2 OR STY0163.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-LT2 / SCS1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McTelland M., Sanderson K.E., Splerth J., Clifton S.W., Latteille P.,
 RA Courtney L., Porwollik S., Ali U., Dante M., Du P., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2.";
 RL Nature 413:852-856(2001).
 CC -1- FUNCTION: Catalyzes the NAD-dependent oxidation of 4-
 CC (phosphohydroxy)-L-threonine (HTP) into 2-amino-3-oxo-4-
 CC (phosphohydroxy)butyric acid which spontaneously decarboxylate to
 CC form 1-amino-3-(phosphohydroxy)propan-2-one (3-amino-2-oxopropyl
 CC phosphate) (By similarity).
 CC -1- CATALYTIC ACTIVITY: 4-(phosphonoxy)-threonine + NAD(+) = 2-
 CC amino-3-oxo-4-phosphonoxybutyrate + NADH.
 CC -1- PATHWAY: De novo synthesis of pyridoxine (vitamin B6) and
 CC pyridoxal phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO THE PDXA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE008701; ALU19127.1; -;
 DR StyGene: SG77777; pdx2.
 DR InterPro: IPR005255; PdxA.
 DR TIGRFAMs: TIGR00557; pdxa: 1.
 KW Pyridoxine biosynthesis; Oxidoreductase; NAD: Complete proteome.
 SQ SEQUENCE 327 AA; 35064 MW; 03BB6725F1896440 CRC64;
 QY 1 RVALAGV 7
 Db 204 RIAVAGV 210
 RESULT 11
 PTDA_ECOLI STANDARD; PRT; 485 AA.
 AC P24241; Q46880;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pdx system, arbutin-, cellobiose-, and salicin-specific IIAAC
 DE component (ETIABC-ASC) (Arbutin-, cellobiose-, and salicin-permease
 DE IIAAC component) (Phosphotransferase enzyme II, ABC component)
 DE (EC 2.7.1.69) (ETI-ASC).
 GN ASCE OR B2715.
 OS Escherichia coli.
 OC Escherichia: Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=92334140; PubMed=1630307;
 RA Hall B.G., Xu L.;
 RT "Nucleotide sequence, function, activation, and evolution of the
 RT cryptic asc operon of *Escherichia coli* K12.";
 RL Mol. Biol. Evol. 9:688-706(1992).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatter F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
 CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS). A MAJOR CARBOHYDRATE ACTIVE
 CC -TRANSPORT SYSTEM. THE IIIC DOMAINS CONTAIN THE SUGAR BINDING SITE
 CC AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY
 CC PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HPR); IIA TRANSFERS ITS
 CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
 CC THE SUGAR.
 CC -1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
 CC histidine + sugar phosphate.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: CONTAINS 1 PTS EIIA DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PTS EIIB DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PTS EIIC DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M73326; AAA16429.1; -;
 DR EMBL: U29579; AAA69225.1; -;
 DR EMBL: AE000355; AAC75757.1; -;
 DR PIR: B44070; B44070.
 DR PIR: S27552; S27552.
 DR HSP: P05053; IIBA.
 DR Ecogene: EG10086; ascf.
 DR InterPro: IPR004719; PTSIIC_glc.
 DR InterPro: IPR001966; PTS_EIIB.
 DR InterPro: IPR003352; PTS_EIIC.
 DR Pfam: PF00367; PTS_EIIB; 1.
 DR Pfam: PF02378; PTS_EIIC; 1.
 DR ProDom: PD001476; PTS_EIIB; 1.
 DR TIGRFAMs: TIGR00826; EIIBC; 1.
 DR TIGRFAMs: TIGR00852; pts-glc; 1.
 DR PROSITE: PS01035; PTS_EIIB_CXS; 1.
 KW Phosphotransferase system; Sugar transport; Transferase;
 KW phosphorylation; Transmembrane; Inner membrane; Complete proteome.
 FT DOMAIN 1 43
 FT DOMAIN ? 485
 FT MOD_RES 28 28
 FT MOD_RES 310 310
 FT TRANSMEM 102 122
 FT TRANSMEM 146 166
 FT TRANSMEM 177 197
 FT TRANSMEM 207 227
 FT TRANSMEM 254 274
 FT TRANSMEM 285 305
 FT TRANSMEM 330 350
 FT TRANSMEM 363 383
 FT TRANSMEM 389 409
 FT TRANSMEM 433 453
 FT CONFLICT 167 170
 FT CONFLICT 311 311
 SQ SEQUENCE 485 AA; 51229 MW; CFEEDDBECAB8C8 CRC64;
 QY 2 VALAGV 8
 Query Match 82.9%; Score 29; DB 1; Length 485;
 Best Local Similarity 85.7%; Pred. No. 1;e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 181 IAIAGVL 187

RESULT 12

ID	NUON_MYCTU	STANDARD	PRT	531 AA
AC	O5308;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	NADH dehydrogenase I chain N (EC 1.6.5.3) (NADH-ubiquinone			
DE	oxidoreductase chain N).			
GN	NUON OR RV3158 OR MT3246 OR MTV014.02C.			
OS	Mycobacterium tuberculosis.			
OC	Bacteria: Actinobacteria: Actinobacteria (class): Actinobacteridae:			
OC	Actinomycetales: Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
OX	NCBI_TaxID=1773;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=H37RV;			
RX	MEDLINE=98295987; PubMed=9634230;			
RA	Cole S.T., Brosch R., Parish J., Garnier T., Churcher C., Harris D.,			
RA	Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,			
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,			
RA	Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,			
RA	Hornsbey T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,			
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,			
RA	Rutler S., Seeger K., Skelton S., Squares S., Squares R.,			
RA	Rutler J.E., Taylor K., Whitehead S., Barrell B.G.;			
RT	"Deciphering the biology of Mycobacterium tuberculosis from the			
RT	complete genome sequence."			
RL	Nature 393:537-544(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CD 1551 / Oshkosh;			
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,			
RA	Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,			
RA	Kolony J.F., Nelson W.C., Umayan L.A., Ermolova M.D., Salzberg S.L.,			
RA	Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,			
RA	Bisbal W.;			
RT	"Whole genome comparison of Mycobacterium tuberculosis clinical and			
RT	laboratory strains."			
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
CC	-1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).			
CC	-1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: AL021646; CA16623.1;			
DR	EMBL: AF007138; AK47585.1;			
DR	TIGR: MT3246;			
DR	TIGR: MT3246;			
DR	TIGR: MT3246;			
DR	InterPro: IPR001750; Oxidored_q1.			
DR	Pfam: PF00361; Oxidored_q1.1.			
KW	Oxidoreductase; NAD; Ubiquinone; Transmembrane; Complete proteome.			
FT	TRANSMEM 8 28			
FT	POTENTIAL.			
FT	TRANSMEM 41 61			
FT	POTENTIAL.			
FT	TRANSMEM 81 101			
FT	POTENTIAL.			
FT	TRANSMEM 146 166			
FT	POTENTIAL.			
FT	TRANSMEM 172 192			
FT	POTENTIAL.			
FT	TRANSMEM 208 228			
FT	POTENTIAL.			
FT	TRANSMEM 250 270			
FT	POTENTIAL.			
FT	TRANSMEM 318 338			
FT	POTENTIAL.			
FT	TRANSMEM 350 370			
FT	POTENTIAL.			
FT	TRANSMEM 372 392			
FT	POTENTIAL.			
FT	TRANSMEM 418 438			
FT	POTENTIAL.			

FT TRANSMEM 453 473 POTENTIAL.

FT TRANSMEM 500 520 POTENTIAL.

FT TRANSMEM 130 130 A -> L (IN REF. 2).

SO SEQUENCE 531 AA; 55342 MW; 45A17C8F3E57BBA1 CRC64;

Query Match 82.9%; Score 29; DB 1; Length 531;

Best Local Similarity 85.7%; Pred. No. 1.2e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 VAIAGVL 8

Db 22 VAAVAVL 28

RESULT 13

ID	LSPA_SYNY3	STANDARD	PRT	161 AA
AC	P73540;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lipoprotein signal peptidase (EC 3.4.23.36) (Lipoprotein signal			
DE	peptidase) (Signal peptidase II) (Spase II).			
GN	LSPA OR SLR1366.			
OS	Synechocystis sp. (strain PCC 6803).			
OC	Bacteria: Cyanobacteria: Chroococcales; Synechocystis.			
OX	NCBI_TaxID=1148;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97061201; PubMed=8905231;			
RA	Kaneke T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,			
RA	Miyajima N., Hiroseawa M., Sugura M., Sasamoto S., Kimura T.,			
RA	Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,			
RA	Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A.,			
RA	Yamada M., Yasuda M., Tabata S.;			
RT	"Sequence analysis of the genome of the unicellular cyanobacterium			
RT	Synechocystis sp. strain PCC6803. II. Sequence determination of the			
RT	entire genome and assignment of potential protein-coding regions."			
RL	DNA Res. 3:109-136(1996).			
CC	-1- FUNCTION: THIS PROTEIN SPECIFICALLY CATALYZES THE REMOVAL OF			
CC	SIGNAL PEPTIDES FROM LIPOPROTEINS (BY SIMILARITY).			
CC	-1- CATALYTIC ACTIVITY: Cleavage of N-terminal leader sequences from			
CC	membrane lipoproteins. Hydrolyzes xaa-xbb-xcc-l-cys, in which			
CC	xaa is hydrophobic (preferably leu), xbb is often ser or ala, xcc			
CC	is often gly or ala, and the cys is alkylated on sulfur with a			
CC	diacylglyceryl group.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A8.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: D90907; BAA17580.1;			
DR	MEROPS: A08.001;			
DR	InterPro: IPR001872; SigPase_A8.			
DR	Pfam: PF01252; Peptidase_A8.1.			
DR	PRINTS: PR00781; LiposigPase.			
DR	ProDom: PD004304; SigPase_A8.1.			
DR	TIGR: TIGR00077; LSPA.1.			
DR	PROSITE: PS00855; SPASE-II.1.			
KW	Hydrolase; Aspartyl protease; Transmembrane; Complete proteome.			
FT	TRANSMEM 11 31			
FT	POTENTIAL.			
FT	TRANSMEM 44 64			
FT	POTENTIAL.			
FT	TRANSMEM 66 86			
FT	POTENTIAL.			
FT	TRANSMEM 135 155			
FT	POTENTIAL.			
FT	ACT_SITE 112 112			
FT	BY SIMILARITY.			
FT	ACT_SITE 137 137			
FT	BY SIMILARITY.			
SO	SEQUENCE 161 AA; 17623 MW; 28D552A10D8FC1DF CRC64;			

Query Match 80.0%; Score 28; DB 1; Length 161;
 Best Local Similarity 62.5%; Pred. No. 66;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVALAGVL 8
 Db 15 QVALAGII 22

RESULT 14

PYRE_COLGR STANDARD; PRT; 233 AA.
 AC P35786;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Orotate phosphoribosyltransferase (EC 2.4.2.10) (OPRT) (OPRTase).
 GN PYR1.
 OS Colletotrichum graminicola (Anthracnose fungus) (Glomerella graminicola).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
 OC Glomerella.
 OX NCBI_TaxID=31870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93062810; PubMed=1435732;
 RA Rasmussen J.B., Panaccione D.G., Pang G.C., Hanau R.M.;
 RT "The PYR1 gene of the plant pathogenic fungus Colletotrichum graminicola: selection by intraspecific complementation and sequence analysis";
 RT Mol. Gen. Genet. 235:74-80(1992).
 CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate + diphosphate = orotate + 5-phospho-alpha-D-ribose 1-diphosphate.
 CC -1- PATHWAY: Pyrimidine biosynthesis; fifth step.
 CC -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE PHOSPHORIBOSYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: S47907; AAB24061.1; -
 CC DR PIR: S30118; S30118.
 CC DR HSP: P08870; 1STO.
 CC DR InterPro: IPR000836; PRTtransferase.
 CC DR InterPro: IPR002375; Pr/PY_TP_transf.
 CC DR Pfam: PF00156; Prbosyltran; 1.
 CC DR TIGRFAMs: TIGR00336; PYRE.1.
 CC DR PROSITE: PS00103; PUR_PYE_PRT_TRANSFER.1.
 CC KW Pyrimidine biosynthesis; Transferase; Glycosyltransferase.
 CC FT ACT_SITE 113 113 BY SIMILARITY.
 CC FT SEQUENCE 233 AA; 25231 MW; 1E6CDB822CA29664 CRC64;

Query Match 80.0%; Score 28; DB 1; Length 233;
 Best Local Similarity 75.0%; Pred. No. 92;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVALAGVL 8
 Db 14 KVALAGVL 21

RESULT 15
 FLIR_TREPA STANDARD; PRT; 265 AA.
 AC P74932;

DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Flagellar biosynthetic protein flir.
 GN FLIR OR TP0716.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_TaxID=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RX MEDLINE=96105201; PubMed=8529894;
 RA Hardham J.M., Frye J.G., Stamm L.V.;
 RT "Identification and sequences of the Treponema pallidum flir, flir, flir and flirB' genes.";
 RL Gene 166:57-64(1995).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RX MEDLINE=98332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
 RA McDonald L., Artlich P., Bowman C., Colton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syphilis spirochete";
 RT Science 281:375-388(1998).
 CC -1- FUNCTION: ROLE IN FLAGELLAR BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE FLIR/MOE/SPAR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U36839; AAB00548.1; -
 CC DR EMBL: AE001244; AAC65682.1; -
 CC DR TIGR: TP0716; -
 CC DR InterPro: IPR002010; Bac_export_1.
 CC DR Pfam: PF01311; Bac_export_1; 1.
 CC DR PRINTS: PR00953; TYPE3IMPRPR.
 CC KW Flagella; Transmembrane; Complete proteome.
 CC FT TRANSMEM 14 34 POTENTIAL.
 CC FT TRANSMEM 44 64 POTENTIAL.
 CC FT TRANSMEM 73 93 POTENTIAL.
 CC FT TRANSMEM 132 152 POTENTIAL.
 CC FT TRANSMEM 182 202 POTENTIAL.
 CC FT TRANSMEM 229 249 POTENTIAL.
 CC FT SEQUENCE 265 AA; 29313 MW; A9BFA16FE8365C09 CRC64;

Query Match 80.0%; Score 28; DB 1; Length 265;
 Best Local Similarity 62.5%; Pred. No. 1e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVALAGVL 8
 Db 42 RVALAGLI 49

Search completed: July 18, 2003, 16:02:20
 Job time : 9.8421 secs

KW ATP-binding; Complete proteome.
SQ SEQUENCE 279 AA; 30525 MW; E051E473D082AD32 CRC64;

Query Match 97.1%; Score 34; DB 16; Length 279;
Best Local Similarity 87.5%; Pred. No. 29;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVAIAGVL 8
Db 153 RVAIAGIL 160

RESULT 14
Q58488 PRELIMINARY; PRT; 279 AA.
AC 058488: 01-JUN-1998 (TREMBLrel. 05, Created)
DT 01-JUN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
GN Cobalt transport ATP-binding protein Cbio.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OX Methanococcaceae; Methanococcus.
RN NCBI_TaxID=2190;
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=66337999; PubMed=868087;
RA Bolt C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhman J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Botodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii."
RL Science 273:1058-1073(1996).
CC -!- FUNCTION: MAY BE INVOLVED WITH COBALT TRANSPORT IN ASSOCIATION WITH COBALAMIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS).
DR EMBL: U67551; AAB99089.1; -.
DR HSSP: Q58663; 1G6H.
DR TIGR: M01088; -.
DR InterPro: IPR003593; AAA_Arpase.
DR InterPro: IPR003439; ABC_transport.
DR Pfam: PF00005; ABC_tran; 1.
DR ProDom: PD000006; ABC_transport; 1.
DR SMART: SM00382; AAA; 1.
DR TIGRfams: TIGR01166; Cbio; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW Cobalamin biosynthesis; Cobalt transport; Transport; Inner membrane;
KW ATP-binding; Complete proteome.
FT NP BIND 37
FT BIND 43
FT SEQUENCE 279 AA; 31120 MW; 7A2DF31BE68CFA81 CRC64;

Query Match 97.1%; Score 34; DB 17; Length 279;
Best Local Similarity 87.5%; Pred. No. 29;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVAIAGVL 8
Db 147 RVAIAGIL 154

RESULT 15
Q09X12 PRELIMINARY; PRT; 280 AA.
AC 09X12;

DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative ABC transporter (ATP-binding protein).
GN SPY2194.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS).
DR EMBL: AE006636; AAK34822.1; -.
DR HSSP: Q58663; 1G6H.
DR InterPro: IPR003593; AAA_Arpase.
DR InterPro: IPR003439; ABC_transport.
DR Pfam: PF00005; ABC_tran; 1.
DR ProDom: PD000006; ABC_transport; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport; Complete proteome.
SQ SEQUENCE 280 AA; 30863 MW; DBE44FEF8780C7D5 CRC64;

Query Match 97.1%; Score 34; DB 16; Length 280;
Best Local Similarity 87.5%; Pred. No. 29;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 RVAIAGVL 8
Db 153 RVAIAGIL 160

Search completed: July 18, 2003, 16:03:46
Job time : 33 secs

SO SEQUENCE 288 AA; 31971 MW; 862E839C8CADEED0 CRC64;
 Query Match 100.0%; Score 35; DB 16; Length 288;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
 Db 153 RVAIAGVL 160
 |||||

RESULT 11
 ID 050801 PRELIMINARY; PRT; 277 AA.
 AC 050801;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Cobalt transport ATP-binding protein CBIO.
 GN CBIO.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97284497; PubMed=9139911;
 RA Eberhardt S., Korn S., Lottspeich F., Bacher A.;
 RL J. Bacteriol. 179:2938-2943(1997).
 CC -1- FUNCTION: MAY BE INVOLVED WITH COBALT TRANSPORT IN ASSOCIATION
 CC WITH COBALAMIN BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS).
 DR EMBL: X94292; CAA63958.1; -
 DR HSPR: Q58663; 166H.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transportr.
 DR Pfam: PF00005; ABC_tran; 1.
 DR Prodom: PD000006; ABC_transportr; 1.
 DR SMART: SM00382; AAA; 1.
 DR TIGRfams: TIGR01166; Cbio. 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KM Cobalamin biosynthesis; Cobalt transport; Transport; Inner membrane;
 FT ATP-binding.
 FT NP_BIND 37 44 ATP (POTENTIAL).
 SQ SEQUENCE 277 AA; 30425 MW; 9B93B1CB042E0D2 CRC64;

Query Match 97.1%; Score 34; DB 1; Length 277;
 Best Local Similarity 87.5%; Pred. No. 29;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
 Db 146 RVAIAGVL 153
 |||||

RESULT 12
 ID 0929J3 PRELIMINARY; PRT; 278 AA.
 AC 0929J3;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE YXA protein (ABC transporter) (ATP-binding protein).
 GN YXA OR BH0164.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;

RX MEDLINE=99209008; PubMed=10192928;
 RA Takami H., Takaki Y., Nakasone K., Hirama C., Inoue A., Horikoshi K.;
 RT "Sequence analysis of a 32-kb region including the major ribosomal
 RT protein gene clusters from alkaliphilic Bacillus sp. strain C-125.";
 RL Biosci. Biotechnol. Biochem. 63:452-455(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS).
 DR EMBL: AB017508; BAA75300.1; -
 DR EMBL: AP001507; BAB03883.1; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transportr.
 DR Pfam: PF00005; ABC_tran; 1.
 DR Prodom: PD000006; ABC_transportr; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KM ATP-binding; Transport; Complete proteome.
 SQ SEQUENCE 278 AA; 30716 MW; BF5FBA2216074698 CRC64;

Query Match 97.1%; Score 34; DB 16; Length 278;
 Best Local Similarity 87.5%; Pred. No. 29;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
 Db 148 RVAIAGVL 155
 |||||

RESULT 13
 ID 097NS1 PRELIMINARY; PRT; 279 AA.
 AC 097NS1;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE ABC transporter, ATP-binding protein.
 GN SP2220.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapple E., Khouri H., Wolf A.M., Urdetack T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae.";
 RL Science 293:498-506(2001).
 DR EMBL: AE007509; AAK76268.1; -
 DR TIGR: SP2220;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transportr.
 DR Pfam: PF00005; ABC_tran; 1.
 DR Prodom: PD000006; ABC_transportr; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.

```

DR InterPro: IPR003439; ABC_transportr.
DR Pfam: PF00005; ABC_tran; 1.
DR ProDom: PD000006; ABC_transportr; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport; Complete proteome.
SQ SEQUENCE 282 AA; 31060 MW; 6D54908503DE47E2 CRC64;

Query Match 100.0%; Score 35; DB 16; Length 282;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
Db 141 RVAIAGVL 148

RESULT 8
O8XNV7 PRELIMINARY; PRT; 285 AA.
ID O8XNV7
AC O8XNV7;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Probable cobalt transport ATP-binding protein O.
GN CBO OR CPE0195.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL: AP003185; BAB79901.1; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transportr.
DR Pfam: PF00005; ABC_tran; 1.
DR ProDom: PD000006; ABC_transportr; 1.
DR SMART: SM00382; AAA; 1.
DR TIGRFAMs: TIGR01166; cblO; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 285 AA; 32133 MW; DB3AE8383CAC5059 CRC64;

Query Match 100.0%; Score 35; DB 16; Length 285;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
Db 149 RVAIAGVL 156

RESULT 9
O927N9 PRELIMINARY; PRT; 288 AA.
ID O927N9
AC O927N9;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical protein lin249.
GN LIN249.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kneft J., Kuhn M., Kunst F., Kurapkut G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Negjari H.,
RA Nordstok G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL: AL596173; CAC97975.1; -.
DR ListList: LIN02749; -.
DR InterPro: IPR003439; ABC_transportr.
DR Pfam: PF00005; ABC_tran; 1.
DR ProDom: PD000006; ABC_transportr; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 288 AA; 31882 MW; F08B7CB20AEC3F9 CRC64;

Query Match 100.0%; Score 35; DB 16; Length 288;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
Db 153 RVAIAGVL 160

RESULT 10
O8Y455 PRELIMINARY; PRT; 288 AA.
ID O8Y455
AC O8Y455;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical protein lmo2600.
GN lmo2600.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-E / SEROVAR 1/2A;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kneft J., Kuhn M., Kunst F., Kurapkut G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Negjari H.,
RA Nordstok G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL: AL591983; CAD00678.1; -.
DR ListList: LMO02600; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transportr.
DR Pfam: PF00005; ABC_tran; 1.
DR ProDom: PD000006; ABC_transportr; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW Hypothetical protein; Complete proteome.

```

RA Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 RT "Comparative genomics of *Listeria* species.";
 RL Science 294:849-852(2001).
 DR EMBL: AL596173; CAC97976.1; -
 DR Listlist; LINO2750; -
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00005; ABC_tran; 1.
 DR ProDom: PD000006; ABC_transport; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 279 AA; 30840 MW; 89E3FB0520FC9391 CRC64;

Query Match 100.0%; Score 35; DB 16; Length 279;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVAIAGVL 8
 |||||
 Db 148 RVAIAGVL 155

RESULT 5
 Q8Y454 PRELIMINARY; PRT; 279 AA.
 ID 08Y454;
 AC 08Y454;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein lmo2601.
 GN LMO2601.
 OS *Listeria monocytogenes*.
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 NCBI_TaxId=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EGD-E / SEROVAR 1/2A;
 RX MEDLINE-21537279; PubMed-11679669;
 RA Clauser P., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Chardit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussangot O.,
 RA Eutian K.-D., Eschl H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkai G.,
 RA Madueno E., Maltounam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 RT "Comparative genomics of *Listeria* species.";
 RL Science 294:849-852(2001).
 DR EMBL: AL591983; CAD00679.1; -
 DR Listlist; LMO2601; -
 DR InterPro: IPR003593; AAA_Arpase.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00005; ABC_tran; 1.
 DR ProDom: PD000006; ABC_transport; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 279 AA; 30911 MW; 8E949071C9EE0791 CRC64;

Query Match 100.0%; Score 35; DB 16; Length 279;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVAIAGVL 8
 |||||
 Db 148 RVAIAGVL 155

RESULT 6
 Q97EK8 PRELIMINARY; PRT; 281 AA.
 ID 097EK8;
 AC 097EK8;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE ABC-type transporter, Arpase component (cobalt transporters
 subfamily).
 GN CAC3102.
 OS *Clostridium acetobutylicum*.
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 NCBI_TaxId=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE-21359325; PubMed-11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hiti J., Wolf Y.I.,
 RA Tatunov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.,
 RT "Genome sequence and comparative analysis of the solvent-producing
 bacterium *Clostridium acetobutylicum*.";
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL: AE007806; AKK81042.1; -
 DR InterPro: IPR003593; AAA_Arpase.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001230; Prenyl_site.
 DR Pfam: PF00005; ABC_tran; 1.
 DR ProDom: PD000006; ABC_transport; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.
 KW Complete proteome.
 SO SEQUENCE 281 AA; 31504 MW; 3864E28BF9A0B42 CRC64;

Query Match 100.0%; Score 35; DB 16; Length 281;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVAIAGVL 8
 |||||
 Db 153 RVAIAGVL 160

RESULT 7
 Q9KGD6 PRELIMINARY; PRT; 282 AA.
 ID 09KGD6;
 AC 09KGD6;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE ABC transporter (ATP-binding protein).
 GN BH0165.
 OS *Bacillus halodurans*.
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 NCBI_TaxId=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RX MEDLINE-20512582; PubMed-11058132;
 RA Takami H., Nakasone K., Takaki Y., Mieno G., Sasaki R., Masui N.,
 RA Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuhara S.,
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
halodurans and genomic sequence comparison with *Bacillus subtilis*.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -I- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 (ABC TRANSPORTERS).
 DR EMBL: AP001507; BAB03884.1; -
 DR InterPro: IPR003593; AAA_Arpase.

AC Q929J2;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE YBAE protein.
 GN YBAE.
 OS *Bacillus halodurans*.
 OC Bacteria; Firmicutes; *Bacillus*/Clostridium group; Bacillales;
 OC Bacillaceae; *Bacillus*.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125;
 RX MEDLINE=9209008; PubMed=10192928;
 RA Takaki Y., Nakasone K., Hirama C., Takami H., Inoue A., Horikoshi K.;
 RT "Sequence analysis of a 32-kb region including the major ribosomal
 RT protein gene clusters from alkaliphilic *Bacillus* sp. strain C-125.";
 RL Biosci. Biotechnol. Biochem. 63:452-455(1999).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: AB017508; BAA75301.1; -;
 DR InterPro: IPR003439; ABC_tran. 1;
 DR Pfam: PF00005; ABC_tran. 1;
 DR Prodom: PD000006; ABC_transporter. 1;
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transport.
 SQ SEQUENCE 183 AA; 20247 MW; 75480181377C2A3F CRC64;

Query Match 100.0%; Score 35; DB 2; Length 183;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
 Db 42 RVAIAGVL 49

RESULT 3
 ID P70970 PRELIMINARY; PRT; 276 AA.
 AC P70970;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein ybae.
 GN YBAE.
 OS *Bacillus subtilis*.
 OC Bacteria; Firmicutes; *Bacillus*/Clostridium group; Bacillales;
 OC Bacillaceae; *Bacillus*.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124188; PubMed=8696501;
 RA Yasumoto K., Iiu H., Jeong S.M., Ohashi Y., Kakihama S., Tanaka K.,
 RA Kawamura F., Yoshikawa H., Takahashi H.;
 RT "Sequence analysis of a 50 kb region between spoOH and trnH on the
 RT *Bacillus subtilis* chromosome.";
 RL Microbiology 142:3039-3046(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Aaevedo V., Bertolo M.G., Bessieres P., Bolotin A., Borcherdt S.,
 RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruchi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings S.J., Daniel R.A.,
 RA Danilov F., Devine K.M., Dusterhoft A., Enrich S.D., Emmerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kunano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Portwollik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Sefor S.J., Sefor P., Shin B.S., Solo B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vanier F., Vassart A.,
 RA Varai A., Wambuit R., Wedler E., Wedler H., Weitzengger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yeta K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RL Nature 390:249-256(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBD databases.
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS).
 DR EMBL: D64126; BAA10984.1; -;
 DR EMBL: Z99104; CAB11922.1; -;
 DR HSP: Q58663; 1GAU.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00005; ABC_tran. 1;
 DR Prodom: PD000006; ABC_transporter. 1;
 DR SMART: SM00382; AAA. 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Hypothetical protein; Transport; Complete proteome.
 SQ SEQUENCE 276 AA; 30578 MW; A5AB1B64360016C9 CRC64;

Query Match 100.0%; Score 35; DB 16; Length 276;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
 Db 140 RVAIAGVL 147

RESULT 4
 ID Q927N8 PRELIMINARY; PRT; 279 AA.
 AC Q927N8;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein lin2750.
 GN LIN2750.
 OS *Listeria innocua*.
 OC Bacteria; Firmicutes; *Bacillus*/Clostridium group; Bacillales;
 OC Listeriaceae; *Listeria*.
 OX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIP 11262 / SEROVAR 6A;
 RX PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deloux P.,
 RA Domann K.-D., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
 RA Entian K.-D., Essli H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
 RA Madueno E., Maltournam A., Mata Vicente J., Ng E., Nedjari H.,

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:59:04 ; Search time 32 seconds
(without alignments)
51.512 Million cell updates/sec

Title: US-10-054-968-8

Perfect score: 35
Sequence: 1 RVAINGVL 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	177	2	Q8VNL8
2	35	100.0	183	2	Q929J2
3	35	100.0	276	16	P70970
4	35	100.0	279	16	Q927N8
5	35	100.0	279	16	Q8Y454
6	35	100.0	281	16	Q97EK8
7	35	100.0	282	16	Q9KGP6
8	35	100.0	285	16	Q8XN7
9	35	100.0	288	16	Q927N9
10	35	100.0	288	16	Q8Y455
11	34	97.1	277	1	Q50801
12	34	97.1	278	16	Q929J3
13	34	97.1	279	16	Q87N51
14	34	97.1	279	17	Q58488
15	34	97.1	280	16	Q99X12
16	34	97.1	286	16	Q8R7Y5

17	34	97.1	288	16	Q9C1S8	Q9C1S8 lactococcus
18	34	97.1	311	17	Q27739	Q27739 methanobact
19	34	97.1	433	16	Q9PPV2	Q9PPV2 ureaplasma
20	33	94.3	150	17	Q9HSP8	Q9HSP8 halobacteri
21	33	94.3	271	16	Q8RHK9	Q8RHK9 fusobacteri
22	33	94.3	281	16	Q8XHV2	Q8XHV2 clostridium
23	33	94.3	285	16	Q8XHV3	Q8XHV3 clostridium
24	33	94.3	453	17	Q8TK65	Q8TK65 methanosarc
25	32	91.4	275	16	Q97N50	Q97N50 streptococc
26	32	91.4	279	2	Q8VNL9	Q8VNL9 enterococcu
27	32	91.4	280	17	Q26236	Q26236 methanobact
28	32	91.4	286	16	Q97EK9	Q97EK9 clostridium
29	32	91.4	293	16	Q8R7Y4	Q8R7Y4 thermoaer
30	32	91.4	304	16	Q980H4	Q980H4 mycoplasma
31	32	91.4	835	5	Q9V744	Q9V744 drosophila
32	32	88.6	180	2	Q56804	Q56804 xanthomonas
33	31	88.6	254	10	Q94EB6	Q94EB6 oryza sativ
34	31	88.6	268	16	Q92CK1	Q92CK1 listeria in
35	31	88.6	268	16	Q8Y7R4	Q8Y7R4 listeria mo
36	31	88.6	271	16	Q825N5	Q825N5 salmonella
37	31	88.6	277	16	Q9C1S9	Q9C1S9 lactococcus
38	31	88.6	287	16	P96668	P96668 bacillus su
39	31	88.6	304	2	Q9RBA5	Q9RBA5 xanthomonas
40	31	88.6	344	16	Q9RVG6	Q9RVG6 delinococcus
41	31	88.6	364	16	Q67641	Q67641 aquilex aeo
42	31	88.6	586	17	Q30231	Q30231 archaeoglob
43	31	88.6	630	17	Q97UR4	Q97UR4 sulfobolus
44	31	88.6	936	5	Q9N9T2	Q9N9T2 leishmania
45	30	85.7	62	17	Q9H539	Q9H539 halobacteri

ALIGNMENTS

RESULT 1	Q8VNL8	PRELIMINARY;	PRT;	177 AA.
ID	Q8VNL8			
AC	Q8VNL8			
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Putative ABC transporter (Fragment).			
OS	Enterococcus faecium (Streptococcus faecium).			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;			
OX	NCBI_TaxID=1352;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Burnie J.P., Carter T.L., Riggs G.P., Hodgetts S.J., Donohoe M.S.,			
RA	Mathews R.C.;			
RT	*Identification of ABC transporter homologs in vancomycin resistant			
RT	Enterococcus faecium as potential targets for antibody therapy.*;			
RL	Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AJ428868; CAD21831.1; -			
DR	InterPro; IPR003439; ABC_transporter.			
DR	Pfam; PF00005; ABC_tran; 1.			
DR	PRODom; PD00006; ABC_transporter; 1.			
DR	PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.			
FT	NON_TER			
SQ	SEQUENCE 177 AA; 19663 MW; B00F2778F106845F CRC64;			
Query Match	100.0%; Score 35; DB 2; Length 177;			
Best Local Similarity	100.0%; Pred. No. 11;			
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 RVAINGVL 8			
DB	153 RVAINGVL 160			
RESULT 2	Q929J2	PRELIMINARY;	PRT;	183 AA.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:57:45 ; Search time 6.63158 Seconds
(Without alignments)
37.526 Million cell updates/sec

Title: US-10-054-968-7
Perfect score: 36
Sequence: 1 FGPKNF 6

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	83.3	76	1 CRP1_HUMAN	P50238 homo sapien
2	30	83.3	76	1 CRP1_MOUSE	P04006 mus musculu
3	30	83.3	132	1 IL4_HORSE	P42202 equus cabal
4	30	83.3	191	1 CYSR_CHICK	P32965 gallus gall
5	30	83.3	192	1 CSR2_HUMAN	O16527 homo sapien
6	30	83.3	192	1 CSR2_RAT	Q62908 rattus norv
7	30	83.3	192	1 CYSR_HUMAN	P21291 homo sapien
8	30	83.3	192	1 CYSR_RAT	P47875 rattus norv
9	30	83.3	193	1 CSR2_CHICK	P50460 gallus gall
10	30	83.3	268	1 CSR2_COTUA	O05158 coturnix co
11	30	83.3	270	1 IF2B_ARATH	O44473 triticum ae
12	30	83.3	271	1 IF2B_WHEAT	P35871 malus domes
13	30	83.3	301	1 YNT4_YEAST	P40169 saccharomyc
14	30	83.3	309	1 YD22_YEAST	O07651 saccharomyc
15	30	83.3	421	1 P2X5_HUMAN	Q34086 homo sapien
16	30	83.3	421	1 P2X5_MOUSE	Q24400 dtrosophila
17	30	83.3	495	1 MLP2_DROME	Q34086 dtrosophila
18	30	83.3	505	1 XAA_MOUSE	Q34086 dtrosophila
19	30	83.3	692	1 HS78_YEAST	Q34086 dtrosophila
20	30	83.3	727	1 MUTB_PROFR	P16533 propionibac
21	30	83.3	840	1 HS74_HUMAN	P45932 homo sapien
22	30	83.3	841	1 HS74_MOUSE	O61316 mus musculu
23	30	83.3	856	1 HI05_CRIGR	Q60446 cricetus
24	30	83.3	4967	1 RYR2_HUMAN	Q92736 homo sapien
25	30	83.3	4969	1 RYR2_MOUSE	P30957 oryctolagus
26	29	80.6	188	1 DCD_NEIMA	Q91868 neisseria m
27	29	80.6	488	1 ASTD_PSEAE	O50114 pseudomonas
28	29	80.6	492	1 ASTD_ECOLI	P76217 escherichia
29	29	80.6	1181	1 MYSC_DICDI	P42522 dictyostell
30	28	77.8	223	1 CYPH_NECOR	P10235 xenopus lae
31	28	77.8	256	1 CAPA_XENLA	P25229 xenopus lae
32	28	77.8	282	1 CAPA_CAEL	P34685 caenorhabdi
33	28	77.8	284	1 CAZ1_MOUSE	P47753 mus musculu

34	28	77.8	286	1 CAZ1_CHICK	P13127 gallus gall
35	28	77.8	286	1 CAZ1_HUMAN	P52907 homo sapien
36	28	77.8	326	1 Y370_MYCPN	P75230 mycoplasma
37	28	77.8	448	1 Y347_CHLPP	O92816 chlamydia p
38	28	77.8	551	1 YVD3_CAEL	P55114 caenorhabdi
39	28	77.8	552	1 C971_PEA	O43078 pisum sativi
40	28	77.8	567	1 D5BD_CAME	O98hr3 campylobact
41	28	77.8	576	1 C972_SOYBN	O48921 glycine max
42	28	77.8	580	1 C973_ARATH	O23365 arabidopsis
43	28	77.8	1084	1 HGP3_HAETH	P44836 haemophilus
44	28	77.8	4085	1 RPPA_CVH22	O05002 human coron
45	27	75.0	117	1 HV1A_HUMAN	P01742 homo sapien

ALIGNMENTS

RESULT 1
CRP1_HUMAN STANDARD: PRT: 76 AA.
AC P50238: Q13628; Q96J34;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteine-rich protein 1 (Cysteine-rich intestinal protein) (CRIP)
DE (Cysteine-rich heart protein) (hCRHP).
GN CRIP1 OR CRIP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=95091772; PubMed=7999070;
RA Tsui S.K., Yam N.Y., Lee C.Y., Waye M.M.;
RT "Isolation and characterization of a cDNA that codes for a LIM-
containing protein which is developmentally regulated in heart.";
RL Blochem. Biophys. Res. Commun. 205:497-505(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=97271694; PubMed=9126610;
RA Khoo C., Blanchard R.K., Sullivan V.K., Cousins R.J.;
RT "Human cysteine-rich intestinal protein: cDNA cloning and expression
of recombinant protein and identification in human peripheral blood
mononuclear cells.";
RL Protein Expr. Purif. 9:379-387(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: SEEMS TO HAVE A ROLE IN ZINC ABSORPTION AND MAY FUNCTION
AS AN INTRACELLULAR ZINC TRANSPORT PROTEIN.
RL -1- SIMILARITY: CONTAINS 1 LIM DOMAIN. THE LIM DOMAIN BINDS 2
ZINC IONS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U09770; AAA64537.1; -
CC EMBL: U58630; AAB61158.1; -
CC EMBL: BC002738; AAH02738.1; -
CC HSSP: P04006; IIML.
CC Genew: HGNC:2360; CRIP1.
CC MIM: 123875; -
CC InterPro: IPR001781; LIM.

DR Pfam: PF00412; LIM; 1.
 DR ProDom: PD000094; LIM; 1.
 DR SMART: SM00132; LIM; 1.
 DR PROSITE: PS00478; LIM_DOMAIN_1; 1.
 DR PROSITE: PS50023; LIM_DOMAIN_2; 1.
 KW LIM domain; Metal-binding; zinc.
 FT INT_MET 0 0 BY SIMILARITY.
 FT DOMAIN 3 60 LIM.
 FT DOMAIN 61 69 GLY-RICH.
 FT CONFLICT 57 57 A->V (IN REF. 1).
 SQ SEQUENCE 76 AA; 8402 MW; ECC3736447BDF97E CRC64;

Query Match 83.3%; Score 30; DB 1; Length 76;
 Best Local Similarity 83.3%; Pred. No. 7.6;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FGPKNF 6
 |||||
 Db 60 FGPKGF 65

RESULT 2
 CRIP_MOUSE STANDARD; PRT; 76 AA.
 ID CRIP_MOUSE
 AC P04006:
 DT 23-OCT-1986 (Rel. 02, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cysteine-rich protein 1 (Cysteine-rich intestinal protein) (CRIP).
 GN CRIP OR CRIP.
 OS Mus musculus (Mouse), and
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090, 10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-RAT; STRAIN-Sprague-Dawley;
 RX MEDLINE=86205983; PubMed=3085096;
 RA Birkenmeier E.H., Gordon J.I.;
 RT "Developmental regulation of a gene that encodes a cysteine-rich
 RT intestinal protein and maps near the murine immunoglobulin heavy
 RT chain locus";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:2516-2520(1986).
 RN [2]
 RP PARTIAL SEQUENCE OF 1-40, AND ZINC-BINDING..
 RC SPECIES-RAT; STRAIN-Sprague-Dawley; TISSUE=Intestine;
 RX MEDLINE=92052153; PubMed=1946385;
 RA Hempe J.M., Cousins R.J.;
 RT "Cysteine-rich intestinal protein binds zinc during transmembrane zinc
 RT transport";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9671-9674(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-MOUSE;
 RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T.,
 RA Geisler S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,
 RA Schellenberg K., Stepoe M., Tan F., Underwood K., Moore B.,
 RA Theising B., Wylie T., Lennon G., Soares B., Wilson R., Waterston R.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP STRUCTURE BY NMR.
 RX MEDLINE=96211640; PubMed=8632457;
 RA Perez-Alvarado G.C., Kosa J.L., Louis H.A., Beckerle M.C., Winge D.R.,
 RA Summers M.F.;
 RT "Structure of the cysteine-rich intestinal protein, CRIP";
 RL J. Mol. Biol. 257:153-174(1996)
 CC -1- FUNCTION: SEEMS TO HAVE A ROLE IN ZINC ABSORPTION AND MAY FUNCTION
 CC AS AN INTRACELLULAR ZINC TRANSPORT PROTEIN.
 CC -1- DEVELOPMENTAL STAGE: THE CONCENTRATION IN INTESTINAL TISSUES OF
 CC THE mRNA FOR CRIP UNDERGOES AN ABRUPT INCREASE DURING THE ANIMAL'S
 CC TRANSITION FROM SUCKLING TO WEANING.
 CC -1- SIMILARITY: CONTAINS 1 LIM DOMAIN. THE LIM DOMAIN BINDS 2

CC ZINC IONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M13018; AAA37458.1; -;
 DR EMBL: A4189594; -; NOT_ANNOTATED_CDS.
 DR EMBL: W42226; -; NOT_ANNOTATED_CDS.
 DR EMBL: A4183798; -; NOT_ANNOTATED_CDS.
 DR EMBL: A4027688; -; NOT_ANNOTATED_CDS.
 DR EMBL: W17624; -; NOT_ANNOTATED_CDS.
 DR EMBL: A4016512; -; NOT_ANNOTATED_CDS.
 DR PIR: A03270; GYRTI.
 DR PDB: 1IML; 1I-JUL-96.
 DR MGD; MGI:88501; CRIP.
 DR InterPro: IPR001781; LIM.
 DR Pfam: PF00412; LIM; 1.
 DR ProDom: PD000094; LIM; 1.
 DR SMART: SM00132; LIM; 1.
 DR PROSITE: PS00478; LIM_DOMAIN_1; 1.
 DR PROSITE: PS50023; LIM_DOMAIN_2; 1.
 KW LIM domain; Metal-binding; zinc; 3D-structure.
 FT INT_MET 0 0
 FT DOMAIN 3 60 LIM.
 FT DOMAIN 61 69 GLY-RICH.
 SQ SEQUENCE 76 AA; 8419 MW; E469D17A4653179E CRC64;

Query Match 83.3%; Score 30; DB 1; Length 76;
 Best Local Similarity 83.3%; Pred. No. 7.6;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FGPKNF 6
 |||||
 Db 60 FGPKGF 65

RESULT 3
 IL4_HORSE STANDARD; PRT; 132 AA.
 ID IL4_HORSE
 AC P42202; O46393;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Interleukin-4 precursor (IL-4) (B-cell stimulatory factor 1) (BSF-1)
 DE (Lymphocyte stimulatory factor 1).
 GN IL4.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 ON NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94317294; PubMed=8042287;
 RA Vandebergiff E.V., Swiderski C.E., Horohov D.W.;
 RT "Molecular cloning and sequencing of equine interleukin 4";
 RL Vet. Immunol. Immunopathol. 40:379-384(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Schrenzel M.D., Stannard A.A., Daft B., Ferrick D.A.;
 RT "Horse (Equus caballus) interleukin-4 gene";
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: IL-4 PARTICIPATES IN AT LEAST SEVERAL B-CELL ACTIVATION
 CC PROCESSES AS WELL AS OF OTHER CELL TYPES. IT IS A COSTIMULATOR OF
 CC DNA-SYNTHESIS. IT INDUCES THE EXPRESSION OF CLASS II MHC MOLECULES
 CC ON RESTING B-CELLS. IT ENHANCES BOTH SECRETION AND CELL SURFACE
 CC EXPRESSION OF IGE AND IGGL. IT ALSO REGULATES THE EXPRESSION OF
 CC THE LOW AFFINITY FC RECEPTOR FOR IGE (CD23) ON BOTH LYMPHOCYTES
 CC AND MONOCYTES.

CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; L06010; AAA21299.1; -;
 DR EMBL; AF035404; AAB87701.1; -;
 DR HSSP; P05112; ILL.
 DR InterPro: IPR002354; Interleukin_4.
 DR InterPro: IPR001325; Interleukin_4_13.
 DR Pfam; PF00727; IL4; 1.
 DR PRINTS; PR00431; INTERLEUKIN4.
 DR PRODOM; PD004070; Interleukin_4; 1.
 DR SMART; SM00190; IL4_13; 1.
 DR PROSITE; PS00838; INTERLEUKIN_4_13; 1.
 DR Cytokine; Glycoprotein; B-cell activation; Growth factor; Signal.
 KW SIGNAL
 FT SIGNAL 1 23 BY SIMILARITY.
 FT CHAIN 24 132 INTERLEUKIN-4.
 FT DISULFID 24 131 BY SIMILARITY.
 FT CARBOHYD 38 88 BY SIMILARITY.
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 5 5 S -> Y (IN REF. 2).
 FT CONFLICT 25 25 K -> I (IN REF. 2).
 FT CONFLICT 35 35 K -> I (IN REF. 2).
 FT CONFLICT 95 95 N -> NL (IN REF. 2).
 FT CONFLICT 114 114 L -> LK (IN REF. 2).
 FT CONFLICT 125 125 R -> K (IN REF. 2).
 SQ SEQUENCE 132 AA; 14631 MW; 7EB975C5394B2AD0 CRC64;
 Query Match 83.3%; Score 30; DB 1; Length 132;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FGPKN 5
 Db 57 FGPKN 61
 CYSR-CHICK STANDARD; PRT; 191 AA.
 ID CYSR-CHICK
 AC P32965;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cysteine-rich protein 1 (CRP1) (CRP).
 GN CSRP1 OR CSRP.
 OS Gallus gallus (Chicken), and
 OS Coturnix coturnix japonica (Japanese quail).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OC NCB1;TaxID=9031; 93934;
 RN [1]
 RN SEQUENCE FROM N.A.
 RN SPECIES-CHICKEN; TISSUE-Embryonic fibroblast;
 RX MEDLINE-94124603; PubMed-8294495.
 RA Crawford A.W., Pino J.D., Beckerle M.C.;
 RT "Biochemical and molecular characterization of the chicken
 RT cysteine-rich protein, a developmentally regulated LIM-domain protein
 RT that is associated with the actin cytoskeleton.";
 RL J. Cell Biol. 124:117-127(1994).
 RN [2]
 RN SEQUENCE FROM N.A.
 RN SPECIES-C.C.japonica; TISSUE-Embryonic fibroblast;
 RX MEDLINE-96081967; PubMed-7499425;

RA Weiskirchen R., Pino J.D., Macaluso T., Bister K., Beckerle M.C.;
 RT "The cysteine-rich protein family of highly related LIM domain
 RT proteins";
 RL J. Biol. Chem. 270:28946-28954(1995).
 RN [3]
 RN PARTIAL SEQUENCE.
 RN SPECIES-CHICKEN;
 RX MEDLINE-93107157; PubMed-1469049;
 RA Sadler I., Crawford A.W., Michelsen J.W., Beckerle M.C.;
 RT "Zyxin and cCRP: two interactive LIM domain proteins associated with
 RT the cytoskeleton.";
 RL J. Cell Biol. 119:1573-1587(1992).
 RN [4]
 RN ZINC-BINDING.
 RN MEDLINE-93281587; PubMed-8506279;
 RA Michelsen J.W., Schweichel K.L., Beckerle M.C., Winge D.R.;
 RT "The LIM motif defines a specific zinc-binding protein domain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:4404-4408(1993).
 RN [5]
 RN MUTAGENESIS.
 RN MEDLINE-94209279; PubMed-8157637;
 RA Michelsen J.W., Sewell A.K., Louis H.A., Olsen J.I., Davis D.R.,
 RT Winge D.R., Beckerle M.C.;
 RL "Mutational analysis of the metal sites in an LIM domain.";
 RN J. Biol. Chem. 269:11108-11113(1994).
 RN [6]
 RN STRUCTURE BY NMR OF C-TERMINAL LIM DOMAIN.
 RX MEDLINE-95393167; PubMed-7664053;
 RA Perez-Alvarado G.C., Miles C., Michelsen J.W., Louis H.A., Winge D.R.;
 RT "Structure of the carboxy-terminal LIM domain from the cysteine rich
 RT protein CRP";
 RL Nat. Struct. Biol. 1:388-398(1994).
 CC -1- FUNCTION: HEAT STABLE PROTEIN. THAT INTERACTS WITH ZYXIN. MAY BE A
 CC COMPONENT OF A SIGNAL TRANSDUCTION PATHWAY THAT MEDIATES ADHESION-
 CC STIMULATED CHANGES IN GENE EXPRESSION.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: ASSOCIATES WITH THE ACTIN CYTOSKELETON.
 CC -1- TISSUE SPECIFICITY: MOST PROMINENT IN TISSUES THAT ARE ENRICHED IN
 CC SMOOTH MUSCLE CELLS, SUCH AS GIZZARD, STOMACH, AND INTESTINE.
 CC LOWER LEVEL IN THE HEART, NO EXPRESSION IN LIVER, SKELETAL MUSCLE,
 CC OR BRAIN.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION LEVELS INCREASE DRAMATICALLY
 CC DURING SMOOTH MUSCLE MATURATION.
 CC -1- DOMAIN: GLYCINE-RICH REPEATS MEDIATE THE ASSOCIATION WITH THE
 CC ACTIN CYTOSKELETON (PROBABLE).
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X73831; CA52053.1; -;
 DR EMBL; Z28533; CA82187.1; -;
 DR PIR; B44358; B44358.
 DR PIR; C44358; C44358.
 DR PIR; D44358; D44358.
 DR PIR; E44358; E44358.
 DR PIR; S34179; S34179.
 DR PIR; A49648; A49648.
 DR PIR; S38879; S38879.
 DR PDB; 1CTL; 03-JUN-95.
 DR InterPro: IPR001781; LIM.
 DR Pfam; PF00412; LIM; 2.
 DR PRODOM; PD000094; LIM; 2.
 DR SMART; SM00132; LIM; 2.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE; PS00023; LIM_DOMAIN_2; 2.
 KW Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc;

KM 3D-structure. 0
 FT INT_MET 9 60 LIM 1.
 FT DOMAIN 62 77 GLY-RICH.
 FT DOMAIN 63 68 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 117 168 LIM 2.
 FT DOMAIN 174 185 GLY-RICH.
 SQ SEQUENCE 191 AA; BC1A45013C7BDA38 CRC64;
 Query Match 83.3%; Score 30; DB 1; Length 191;
 Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FGPKNF 6
 DB 173 FGPKGF 178
 RESULT 5
 CSR2_HUMAN STANDARD; PRT; 192 AA.
 ID CSR2_HUMAN 016527; 093030;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Smooth muscle cell LIM protein (Cysteine-rich protein 2) (CRP2).
 GN CSRP2 OR SMLIM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aorta;
 RX MEDLINE=96215313; PubMed=8626582;
 RA Jain M., Fujita K.P., Hsieh C.-M., Endege W.O., Sibinga N.E.S.,
 RA Yet S.-F., Kashiki S., Lee W.-S., Perrella M.A., Haber E., Lee M.-E.;
 RT "Molecular cloning and characterization of Smlim, a developmentally
 RT regulated LIM protein preferentially expressed in aortic smooth
 RT muscle cells.";
 RL J. Biol. Chem. 271:10194-10199(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97432823; PubMed=9286703;
 RA Weiskirchen R., Eidel M., Utermann G., Bister K.;
 RT "Cloning, structural analysis, and chromosomal localization of the
 RT human CSRP2 gene encoding the LIM domain protein CRP2.";
 RL Genomics 44:83-93(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: DRASTICALLY DOWN-REGULATED IN RESPONSE TO PGCF-BB OR
 CC CELL INJURY, THAT PROMOTE SMOOTH MUSCLE CELL PROLIFERATION AND
 CC DIFFERENTIATION. SEEMS TO PLAY A ROLE IN THE DEVELOPMENT OF THE
 CC EMBRYONIC VASCULAR SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE AORTA, BUT NOT IN
 CC HEART AND SKELETAL MUSCLE.
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: UA6006; AAC27344.1;
 CC EMBL: U57646; AAC51753.1; -.

DR EMBL: U95018; AAC51755.1; -.
 DR EMBL: BC000992; AAH00992.1; -.
 DR HSSP: Q05158; 10LI.
 DR Genew: HGNC:2470; CSRP2.
 DR MIM: 601871; -.
 DR InterPro: IPR001781; LIM.
 DR Pfam: PF00412; LIM; 2.
 DR ProDom: PD000094; LIM; 2.
 DR SMART: SM00132; LIM; 2.
 DR PROSITE: PS00478; LIM DOMAIN 1; 2.
 DR PROSITE: PS00023; LIM DOMAIN 2; 2.
 DR PROSITE: PS00023; LIM DOMAIN 2; 2.
 KW Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc;
 KW Developmental protein; Differentiation.
 FT INT_MET 0 60
 FT DOMAIN 9 60 LIM 1.
 FT DOMAIN 62 77 GLY-RICH.
 FT DOMAIN 63 68 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 118 169 LIM 2.
 FT DOMAIN 175 186 GLY-RICH.
 SQ SEQUENCE 192 AA; 8EEDF84042DA8981 CRC64;
 Query Match 83.3%; Score 30; DB 1; Length 192;
 Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FGPKNF 6
 DB 174 FGPKGF 179
 RESULT 6
 CSR2_RAT STANDARD; PRT; 192 AA.
 ID CSR2_RAT 062908;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Smooth muscle cell LIM protein (Cysteine-rich protein 2) (CRP2).
 GN CSRP2 OR SMLIM.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Aorta, and Smooth muscle;
 RX MEDLINE=96215313; PubMed=8626582;
 RA Jain M., Fujita K.P., Hsieh C.-M., Endege W.O., Sibinga N.E.S.,
 RA Yet S.-F., Kashiki S., Lee W.-S., Perrella M.A., Haber E., Lee M.-E.;
 RT "Molecular cloning and characterization of Smlim, a developmentally
 RT regulated LIM protein preferentially expressed in aortic smooth
 RT muscle cells.";
 RL J. Biol. Chem. 271:10194-10199(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC CELL INJURY, THAT PROMOTE SMOOTH MUSCLE CELL PROLIFERATION AND
 CC DIFFERENTIATION. SEEMS TO PLAY A ROLE IN THE DEVELOPMENT OF THE
 CC EMBRYONIC VASCULAR SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE AORTA; WEAKLY FOUND IN
 CC THE KIDNEY, THYMUS, AND INTESTINE. BARELY DETECTABLE IN BRAIN,
 CC TESTIS, ESOPHAGUS, LUNG, LIVER, AORTIC ADVENTITIA, VENA CAVA, OR
 CC UTERUS; NOT PRESENT IN HEART AND SKELETAL MUSCLE.
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U44948; AAC52554.1; -.
 DR HSSP: Q05158; 1QLI.
 DR InterPro: IPR001781; LIM.
 DR Pfam: PF00412; LIM; 2.
 DR ProDom: PD000094; LIM; 2.
 DR SMART: SM00132; LIM; 2.
 DR PROSITE: PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE: PS50023; LIM_DOMAIN_2; 2.
 DR Nucleic acid protein: Repeat; LIM domain; Metal-binding; Zinc;
 DR Developmental protein; Differentiation.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT DOMAIN 9 60 LIM 1.
 FT DOMAIN 62 77 GLY-RICH.
 FT DOMAIN 63 68 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 118 169 LIM 2.
 FT DOMAIN 175 186 GLY-RICH.
 SQ SEQUENCE 192 AA; 20809 MW; 74BDE00895B6FDA5 CRC64;

Query Match
 Best Local Similarity 83.3%; Score 30; DB 1; Length 192;
 Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 |||||
 DB 174 FGPKGF 179

RESULT 7

CYSR_HUMAN STANDARD; PRT; 192 AA.

ID CYSR_HUMAN
 AC P21291;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cysteine-rich protein 1 (CRP1) (CRP).
 GN CSRP1 OR CSRP OR CYP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Term placenta;
 RX MEDLINE=90326508; PubMed=2115670;
 RA Liehaber S.A., Emery J.G., Urbanek M., Wang X., Cooke N.E.;
 RT "Characterization of a human CDNA encoding a widely expressed and
 RT highly conserved cysteine-rich protein with an unusual zinc-finger
 RT motif.";
 RL Nucleic Acids Res. 18:3871-3879(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92250516; PubMed=1374386;
 RA Wang X., Lee G., Liehaber S.A., Cooke N.E.;
 RT "Human cysteine-rich protein. A member of the LIM/double-finger
 RT family displaying coordinate serum induction with c-myc.";
 RL J. Biol. Chem. 267:9176-9184(1992).
 CC -1- FUNCTION: COULD PLAY A ROLE IN NEURONAL DEVELOPMENT.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M33146; AAA5720.1; -.
 DR EMBL: M76378; AAA58431.1; -.
 DR EMBL: M76376; AAA58431.1; JOINED.

DR EMBL: M76377; AAA58431.1; JOINED.
 DR PIR: S12658; S12658.
 DR PIR: AA0194; AA0194.
 DR HSSP: P32965; ICTL.
 DR Genew: HGNC:2469; CSRP1.
 DR MIM: 123876; -.
 DR InterPro: IPR001781; LIM.
 DR Pfam: PF00412; LIM; 2.
 DR ProDom: PD000094; LIM; 2.
 DR SMART: SM00132; LIM; 2.
 DR PROSITE: PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE: PS50023; LIM_DOMAIN_2; 2.
 DR Nucleic acid protein: Repeat; LIM domain; Metal-binding; Zinc;
 DR Developmental protein; Differentiation.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT DOMAIN 9 60 LIM 1.
 FT DOMAIN 62 77 GLY-RICH.
 FT DOMAIN 63 68 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 118 169 LIM 2.
 FT DOMAIN 175 186 GLY-RICH.
 SQ SEQUENCE 192 AA; 20436 MW; 0CC6B9B35DC726FC CRC64;

Query Match
 Best Local Similarity 83.3%; Score 30; DB 1; Length 192;
 Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 |||||
 DB 174 FGPKGF 179

RESULT 8

CYSR_RAT STANDARD; PRT; 192 AA.

ID CYSR_RAT
 AC P47875;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cysteine-rich protein 1 (CRP1) (CRP).
 GN CSRP1 OR CSRP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Olfactory bulb;
 RX MEDLINE=95116342; PubMed=7816640;
 RA McLaughlin C.R., Tao O., Apood M.E.;
 RT "Isolation and developmental expression of a rat CDNA encoding a
 RT cysteine-rich zinc finger protein.";
 RL Nucleic Acids Res. 22:5477-5483(1994).
 CC -1- FUNCTION: COULD PLAY A ROLE IN NEURONAL DEVELOPMENT.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U09567; AAC52157.1; -.
 DR HSSP: P32965; ICTL.
 DR InterPro: IPR001781; LIM.
 DR Pfam: PF00412; LIM; 2.
 DR ProDom: PD000094; LIM; 2.
 DR SMART: SM00132; LIM; 2.
 DR PROSITE: PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE: PS50023; LIM_DOMAIN_2; 2.

KW Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT DOMAIN 9 60 LIM 1.
 FT DOMAIN 62 77 GLY-RICH.
 FT DOMAIN 63 68 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 118 169 LIM 2.
 FT DOMAIN 175 186 GLY-RICH.
 SQ SEQUENCE 192 AA; 20482 MW; AZEDD4IA66B901F2 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 192;
 Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FGPKNF 6
 DB 174 FGPKGF 179

RESULT 9

CSR2_CHICK STANDARD; PRT; 193 AA.
 ID CSR2_CHICK
 AC P50460;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cysteine-rich protein 2 (CRP2) (Beta-cysteine-rich protein) (Beta-CRP).
 GN CSR2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic fibroblast;
 RX MEDLINE=96081967; PubMed=7499425.
 RA Weiskirchen R., Pino J.D., Macalima T., Bister K., Beckerle M.C.;
 RT "The cysteine-rich protein family of highly related LIM domain proteins.";
 RL J. Biol. Chem. 270:28946-28954(1995).
 CC -1- FUNCTION: TOTALLY DOWN-REGULATED IN TRANSFORMED CELLS. MAY THEREFORE TAKE PART IN THE CONTROL OF CELL GROWTH AND DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC IONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X84264; CAAS9025.1; -
 CC HSSP: Q05158; IAT1.
 CC InterPro: IPR001781; LIM.
 CC Pfam: PF00412; LIM; 2.
 CC ProDom: PD000094; LIM; 2.
 CC SMART: SM00132; LIM; 2.
 CC PROSITE: PS00478; LIM_DOMAIN_1; 2.
 CC PROSITE: PS50023; LIM_DOMAIN_2; 2.
 CC Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc;
 CC Developmental protein; Differentiation.
 CC INIT_MET 0 0 BY SIMILARITY.
 CC DOMAIN 9 60 LIM 1.
 CC DOMAIN 62 77 GLY-RICH.
 CC DOMAIN 63 68 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC DOMAIN 119 170 LIM 2.
 CC DOMAIN 176 187 GLY-RICH.
 SQ SEQUENCE 193 AA; 20793 MW; 972C62E912D5D82B CRC64;

Query Match 83.3%; Score 30; DB 1; Length 193;
 Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FGPKNF 6
 DB 175 FGPKGF 180

RESULT 10

CSR2_COTUA STANDARD; PRT; 193 AA.
 ID CSR2_COTUA
 AC 005158;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cysteine-rich protein 2 (CRP2).
 GN CSR2.
 OS Coturnix coturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.
 OX NCBI_TaxID=93934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic fibroblast;
 RX MEDLINE=93368932; PubMed=8361751.
 RA Weiskirchen R., Bister K.;
 RT "Suppression in transformed avian fibroblasts of a gene (crp) encoding a cysteine-rich protein containing LIM domains.";
 RL Oncogene 8:2317-2324(1993).
 RN [2]
 RP STRUCTURE BY NMR OF 7-66.
 RX MEDLINE=98254497; PubMed=9585524;
 RA Kontaris G., Konrat R., Kraeutler B., Weiskirchen R., Bister K.;
 RT "Structure and intramolecular dynamics of the amino-terminal LIM domain from quail cysteine- and glycine-rich protein CRP2.";
 RL Biochemistry 37:7127-7134(1998).
 RN [3]
 RP STRUCTURE BY NMR OF 116-174.
 RX MEDLINE=97277343; PubMed=9115265;
 RA Konrat R., Weiskirchen R., Kraeutler B., Bister K.;
 RT "Solution structure of the carboxyl-terminal LIM domain from quail cysteine-rich protein CRP2.";
 RL J. Biol. Chem. 272:12001-12007(1997).
 CC -1- FUNCTION: INTERACTS WITH ZYXIN. MAY BE A COMPONENT OF A SIGNAL TRANSDUCTION PATHWAY THAT MEDIATES ADHESION-STIMULATED CHANGES IN GENE EXPRESSION. TOTALLY DOWN-REGULATED IN TRANSFORMED CELLS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC IONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z21643; CAA9759.1; -
 CC PIR: S31844; S31844.
 CC PDB: 1O1T; 20-AUG-97.
 CC PDB: 1A7I; 27-MAY-98.
 CC InterPro: IPR001781; LIM.
 CC Pfam: PF00412; LIM; 2.
 CC ProDom: PD000094; LIM; 2.
 CC SMART: SM00132; LIM; 2.
 CC PROSITE: PS00478; LIM_DOMAIN_1; 2.
 CC PROSITE: PS50023; LIM_DOMAIN_2; 2.
 CC Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc;
 CC Developmental protein; Differentiation; 3D-structure.
 KW

FT INIT_MET 0 0 BY SIMILARITY.
 FT DOMAIN 9 60 LIM 1.
 FT DOMAIN 62 77 GLY-RICH.
 FT DOMAIN 63 68 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 119 170 LIM 2.
 FT DOMAIN 176 187 GLY-RICH.
 SQ SEQUENCE 193 AA; 20779 MW; 8A9C09EDA7C1896E CRC64;

Query Match 83.3%; Score 30; DB 1; Length 193;
 Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 11111
 Db 175 FGPKNF 180

RESULT 11

IF2B_ARATH STANDARD; PRT; 268 AA.
 AC 041969; GRC5N7;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Eukaryotic translation initiation factor 2 beta subunit (eif-2-beta).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Browning K.S., Chen R.;
 RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.

CC SEQUENCE OF 97-252 FROM N.A.

CC STRAIN=cv. Columbia;

CC Desprez T., Amselem J., Chiapello H., Rouze P., Caboche M.,
 RA Hofte H.;

CC Submitted (NOV-1992) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: EIF-2 FUNCTIONS IN THE EARLY STEPS OF PROTEIN SYNTHESIS
 BY FORMING A TERNARY COMPLEX WITH GTP AND INITIATOR TRNA. THIS
 COMPLEX BINDS TO A 40S RIBOSOMAL SUBUNIT, FOLLOWED BY TRNA BINDING
 TO FORM A 43S PREINITIATION COMPLEX. JUNCTION OF THE 60S RIBOSOMAL
 SUBUNIT TO FORM THE 80S INITIATION COMPLEX IS PRECEDED BY
 HYDROLYSIS OF THE GTP BOUND TO EIF-2 AND RELEASE OF AN EIF-2-GDP
 BINARY COMPLEX. IN ORDER FOR EIF-2 TO RECYCLE AND CATALYZE ANOTHER
 ROUND OF INITIATION, THE GDP BOUND TO EIF-2 MUST EXCHANGE WITH GTP
 BY WAY OF A REACTION CATALYZED BY EIF-2B (BY SIMILARITY).
 CC -1- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA, AND A GAMMA
 CHAIN.

CC -1- SIMILARITY: BELONGS TO THE EIF-2-BETA / EIF-5 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF353095; CAK29672.1; -

CC EMBL; Z18133; GAA79110.1; -

CC InterPro: IPR002735; eif5_eif2b.

CC Pfam: PF01873; eif5_eif2b; 2.

CC ProDom: PD004078; eif5_eif2b; 1.

CC Initiation factor; Protein biosynthesis; Zinc-finger.

CC ZN_FING 222 246 C4-TYPE (POTENTIAL).

CC FT DOMAIN 22 28 POLY-LYS (BASIC).

CC FT DOMAIN 64 69 POLY-LYS (BASIC).

CC SEQUENCE 268 AA; 30633 MW; BA7BD3C257222097 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 268;

Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 11111
 Db 203 FAPKNF 208

RESULT 12

IF2B_WHEAT STANDARD; PRT; 270 AA.
 ID IF2B_WHEAT
 AC 024473;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Eukaryotic translation initiation factor 2 beta subunit (eif-2-beta)
 DE (P38).
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
 OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=97329116; PubMed=9185629;
 RA Metz A.M., Browning K.S.;

RT "Assignment of the beta-subunit of wheat eif2 by protein and DNA
 RT sequence analysis and immunoblotting";
 RL Arch. Biochem. Biophys. 342:187-189(1997).

CC -1- FUNCTION: EIF-2 FUNCTIONS IN THE EARLY STEPS OF PROTEIN SYNTHESIS
 BY FORMING A TERNARY COMPLEX WITH GTP AND INITIATOR TRNA. THIS
 COMPLEX BINDS TO A 40S RIBOSOMAL SUBUNIT, FOLLOWED BY TRNA BINDING
 TO FORM A 43S PREINITIATION COMPLEX. JUNCTION OF THE 60S RIBOSOMAL
 SUBUNIT TO FORM THE 80S INITIATION COMPLEX IS PRECEDED BY
 HYDROLYSIS OF THE GTP BOUND TO EIF-2 AND RELEASE OF AN EIF-2-GDP
 BINARY COMPLEX. IN ORDER FOR EIF-2 TO RECYCLE AND CATALYZE ANOTHER
 ROUND OF INITIATION, THE GDP BOUND TO EIF-2 MUST EXCHANGE WITH GTP
 BY WAY OF A REACTION CATALYZED BY EIF-2B (BY SIMILARITY).
 CC -1- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA, AND A GAMMA
 CHAIN.

CC -1- SIMILARITY: BELONGS TO THE EIF-2-BETA / EIF-5 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U87163; AAB65774.1; -

CC InterPro: IPR002735; eif5_eif2b.

CC Pfam: PF01873; eif5_eif2b; 1.

CC ProDom: PD004078; eif5_eif2b; 1.

CC Initiation factor; Protein biosynthesis; Zinc-finger.

CC FT DOMAIN 23 29 POLY-LYS.

CC FT DOMAIN 63 68 POLY-LYS.

CC SEQUENCE 270 AA; 30389 MW; DIAED918C2F36455 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 270;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 11111
 Db 203 FAPKNF 208

RESULT 13

IF2B_MALDO STANDARD; PRT; 271 AA.
 ID IF2B_MALDO
 AC P55871;
 DT 01-NOV-1997 (Rel. 35, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Eukaryotic translation initiation factor 2 beta subunit (eif-2-beta).
 OS *Malus domestica* (Apple) (*Malus sylvestris*).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eustosids I; Rosales; Rosaceae; Maloideae; *Malus*.
 RX NCBI_TaxID=3750;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Granny Smith;
 RA Dong Y.-H., Janssen B.J., Bielecki L.L., Atkinson R.G.,
 RA Morris B.A., Gardner R.C.;
 RT "Isolating and characterizing genes differentially expressed early in
 RT apple fruit development.";
 RL J. Am. Soc. Hort. Sci. 122:752-757(1997).
 CC -1- FUNCTION: EIF-2 FUNCTIONS IN THE EARLY STEPS OF PROTEIN SYNTHESIS
 CC BY FORMING A TERNARY COMPLEX WITH GTP AND INITIATOR TRNA. THIS
 CC COMPLEX BINDS TO A 40S RIBOSOMAL SUBUNIT, FOLLOWED BY MRNA BINDING
 CC TO FORM A 43S PREINITIATION COMPLEX. JUNCTION OF THE 60S RIBOSOMAL
 CC SUBUNIT TO FORM THE 80S INITIATION COMPLEX IS PRECEDED BY
 CC HYDROLYSIS OF THE GTP BOUND TO EIF-2 AND RELEASE OF AN EIF-2-GDP
 CC BINARY COMPLEX. IN ORDER FOR EIF-2 TO RECYCLE AND CATALYZE ANOTHER
 CC ROUND OF INITIATION, THE GDP BOUND TO EIF-2 MUST EXCHANGE WITH GTP
 CC BY WAY OF A REACTION CATALYZED BY EIF-2B (BY SIMILARITY).
 CC -1- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA, AND A GAMMA
 CC CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE EIF-2-BETA / EIF-5 FAMILY.
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U80269; AAC06384.1; ALT_INIT.
 DR InterPro; IPR002735; eIF5_eIF2B.
 DR Pfam; PF01873; eIF5_eIF2B; 2.
 DR ProDom; PD004078; eIF5_eIF2B; 1.
 KW Initiation factor; Protein biosynthesis; Zinc-finger.
 FT ZN.FING 223 247 C4-TYPE (POTENTIAL).
 FT DOMAIN 24 29 POLY-LYS (BASIC).
 FT DOMAIN 65 70 POLY-LYS (BASIC).
 SQ SEQUENCE 271 AA; 30343 MW; 516C98C17A52BA77 CRC64;
 Query Match 83.3%; Score 30; DB 1; Length 271;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GPKNF 6
 DB 204 FAPKNF 209
 RESULT 14
 YNT4_YEAST
 ID YNT4_YEAST STANDARD; PRT; 301 AA.
 AC P40169;
 DT 01-FEB-1995 (Rel. 31, Last Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 34.1 kDa protein in WHI3-CHS1 intergenic region.
 GN YNL194C OR NI394.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 RX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / FY1679;
 RX MEDLINE=95242839; PubMed=7725799;

RA Jonniaux J.-L., Coster F., Purnelle B., Goffeau A.;
 RT "A 21.7 kb DNA segment on the left arm of yeast chromosome XIV
 RT carries WHI3, GCR2, SPX18, SPX19, an homologue to the heat shock gene
 RT SSI1 and 8 new open reading frames of unknown function.";
 RL Yeast 10:1639-1645(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Obermaier B., Piravandi E., Rinke M., Domsley H.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO YEAST SUR7 AND YDL222C.
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X78898; CA55514.1; -
 DR EMBL; Z71470; CA96088.1; -
 DR SGD; S0005138; YNL194C.
 KW Hypothetical protein; Transmembrane
 FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 6 26 POTENTIAL.
 FT DOMAIN 27 112 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 113 133 POTENTIAL.
 FT DOMAIN 134 143 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 144 164 POTENTIAL.
 FT DOMAIN 165 191 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 192 212 POTENTIAL.
 FT DOMAIN 213 301 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 301 AA; 34081 MW; A62F153F41B36174 CRC64;
 Query Match 83.3%; Score 30; DB 1; Length 301;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GPKNF 6
 DB 186 GPKNF 190
 RESULT 15
 YD22_YEAST
 ID YD22_YEAST STANDARD; PRT; 309 AA.
 AC Q07651;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 34.1 kDa protein in CDC13-GCSI intergenic region.
 GN YDL222C.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 RX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rasmussen S.W.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO YEAST SUR7 AND YNL194C.
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL; Z74270; CAA98801.1; -.
DR SGD; S0002381; YDL222C.
KW Hypothetical protein; Transmembrane.
FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 6 26
FT DOMAIN 27 116 POTENTIAL.
FT TRANSMEM 117 137 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 138 140 POTENTIAL.
FT TRANSMEM 141 161 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 162 188 POTENTIAL.
FT TRANSMEM 189 209 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 210 309 POTENTIAL.
FT CARBOHYD 73 73 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 309 AA; 34135 MW; DC1BD70FA52E7A7 CRC64; N-LINKED (GLCNAC. . .) (POTENTIAL).
Query Match 83.3%; Score 30; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GPKNF 6
Db 185 GPKNF 189

Search completed: July 18, 2003, 16:02:19
Job time : 7.63158 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:59:04 ; Search time 24 Seconds
(without alignments)
51.512 Million cell updates/sec

Title: US-10-054-968-7
Perfect score: 36
Sequence: 1 FGPKNF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	177	2	Q8VNL8 enterococcu
2	36	100.0	286	16	Q99S48 staphylococ
3	33	91.7	283	17	Q8TIL5 methanosarc
4	33	91.7	336	5	Q8SV21 encephalito
5	33	91.7	467	16	Q8U989 agrobacteri
6	33	91.7	798	5	Q8RT21 dictyosteli
7	33	91.7	1518	5	Q86316 schistosoma
8	32	88.9	279	16	Q97N51 streptococ
9	32	88.9	280	16	Q99X12 streptococ
10	32	88.9	288	16	Q9CIS8 lactococcus
11	32	88.9	569	16	Q8R9L8 thermoaer
12	32	88.9	912	10	Q9SUE7 arabidopsis
13	32	88.9	1206	9	Q8W6C3 saccharomon
14	31	86.1	307	10	Q9ZUP7 arabidopsis
15	31	86.1	593	10	Q9C6S0 arabidopsis
16	31	86.1	595	10	Q93VK5 arabidopsis

17	31	86.1	659	10	Q8W131 skeletonema
18	31	86.1	783	16	Q985F3 rhizobium 1
19	31	86.1	1014	11	Q92012 mus musculu
20	31	86.1	1043	11	Q99N32 mus musculu
21	30	83.3	123	2	Q48512 lactobacilli
22	30	83.3	148	4	Q9B0K9 homo sapien
23	30	83.3	168	4	Q9B7A4 homo sapien
24	30	83.3	192	13	Q9YGP7 oncorhynch
25	30	83.3	193	11	Q9D1M8 mus musculu
26	30	83.3	193	11	Q9CZG5 mus musculu
27	30	83.3	193	11	Q97315 mus musculu
28	30	83.3	193	11	P97314 mus musculu
29	30	83.3	208	5	Q76574 caenorhabdi
30	30	83.3	213	16	Q97M44 clostridium
31	30	83.3	215	16	Q9WXZ9 thermotoga
32	30	83.3	224	5	Q8T2H6 dictyosteli
33	30	83.3	237	5	Q9U8N8 drosophila
34	30	83.3	241	5	Q9N2W9 caenorhabdi
35	30	83.3	254	5	Q9TW68 drosophila
36	30	83.3	260	17	Q57872 pyrococcus
37	30	83.3	264	16	Q92KW4 rhizobium m
38	30	83.3	271	5	Q9NAQ4 caenorhabdi
39	30	83.3	276	16	P70970 bacillus su
40	30	83.3	288	5	Q20901 caenorhabdi
41	30	83.3	288	16	Q927N9 listeria in
42	30	83.3	288	16	Q8Y455 listeria mo
43	30	83.3	291	5	Q9NGB8 tripeustes
44	30	83.3	291	5	Q9N305 caenorhabdi
45	30	83.3	297	2	Q9R091 buchnera ap

ALIGNMENTS

RESULT 1

ID Q8VNL8 PRELIMINARY; PRT; 177 AA.
AC Q8VNL8:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative ABC transporter (Fragment).
OS Enterococcus faecium (Streptococcus faecium).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
CC Enterococcaceae; Enterococcus.
OX NCBI TaxID=1352;
RN [1]
RP SEQUENCE FROM N.A.
RA Burnie J.P., Carter T.L., Rigg G.P., Hodgetts S.J., Donohoe M.S.,
RA Matthews R.C.;
RT "Identification of ABC transporter homologs in vancomycin resistant
RT Enterococcus faecium as potential targets for antibody therapy.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ428868; CAD21831.1; .
DR InterPro; IPR003439; ABC transportr.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transportr; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
FT NON_TER 177
SQ SEQUENCE 177 AA; 19663 MW; B00F2778F106845F CRC64;

Query Match 100.0%; Score 36; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. NO. 4.7; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FGPKNF 6
|||
Db 109 FGPKNF 114

RESULT 2
Q99S48 PRELIMINARY; PRT; 286 AA.
ID Q99S48

AC Q99S48;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein SAV2221.
 GN SAV2221 OR SA2020.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OS Staphylococcus aureus (strain N315).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 CC Staphylococcus.
 OX NCBI_TaxID=158878, 158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-O., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kamehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 aureus".
 RL Lancet 357:1225-1240(2001).
 DR EMBL; AP003364; BAB58383.1; -;
 DR EMBL; AP003136; BAB43313.1; -;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transportr.
 DR Pfam: PF00005; ABC_tran; 1.
 DR ProDom; PD000006; ABC_transportr; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
 DR Hypothetical protein; Complete proteome.
 SO SEQUENCE 286 AA; 32919 MW; 83DFD065FF30EC8 CRC64;

Query Match 100.0%; Score 36; DB 16; Length 286;
 Best Local Similarity 100.0%; Pred. No. 7.7;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 FGPKNF 6
 Db 109 FGPKNF 114

RESULT 3
 ID 08T15 PRELIMINARY; PRT; 283 AA.
 AC 08T15;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Membrane associated ATPase.
 GN MA4342.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 CC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atwood D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Kizycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umeyam L.A., White O., White R.H., de Macario E.C.,
 RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity."
 RL Genome Res. 12:532-542(2002).
 DR EMBL; AE011152; AM07685.1; -;
 KW Complete proteome.
 SO SEQUENCE 283 AA; 31670 MW; 42E4AC69C1549D5B CRC64;

Query Match 91.7%; Score 33; DB 17; Length 283;
 Best Local Similarity 83.3%; Pred. No. 34;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 FGPKNF 6
 Db 102 FGPKNF 107

RESULT 4
 ID 08SV21 PRELIMINARY; PRT; 336 AA.
 AC 08SV21;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein ECU07_0580.
 GN ECU07_0580.
 OS Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
 OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
 OX NCBI_TaxID=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RA Genoscope;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RX MEDLINE=21576510; PubMed=11719806;
 RA Kallina M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
 RA Prenster G., Barbe V., Peyretailade E., Brothier P., Winkler P.,
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
 RA Weissenbach J., Vlavres C.P.;
 RT "Genome sequence and gene compaction of the eukaryote parasite
 RT Encephalitozoon cuniculi".
 RL Nature 414:450-453(2001).
 DR EMBL; AL590447; CAD25590.1; -;
 KW Hypothetical protein.
 SO SEQUENCE 336 AA; 38458 MW; 899046FE90C5B75 CRC64;

Query Match 91.7%; Score 33; DB 5; Length 336;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 FGPKNF 6
 Db 251 FGPKNF 256

RESULT 5
 ID 08U989 PRELIMINARY; PRT; 467 AA.
 AC 08U989;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE DNA methyltransferase.
 GN ATU3839 OR AGR_L_1997.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. St.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavon T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*
 C58.";
 RL Science 294:2317-2323(2001).
 RM [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21608551; PubMed-11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Quirillo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
 RA Houmlet K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughly D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT *Agrobacterium tumefaciens* C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009315; AAL44649.1;
 DR EMBL; AE008299; AAK89570.1; ALT_INT.
 KW Methyltransferase; Transferase; Complete proteome.
 SQ SEQUENCE 467 AA; 53298 MW; 46DE12AB86162BDA CRC64;

Query Match 91.7%; Score 33; DB 16; Length 467;
 Best Local Similarity 83.3%; Pred. NO. 57;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 Db 45 FGPKNF 50

RESULT 6
 ID Q872L1 PRELIMINARY; PRT; 798 AA.
 AC Q872L1;
 DT 01-JUN-2002 (TRENBLREL. 21, Created)
 DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE Hypothetical 93.1 kDa protein.
 OS *Dictyostelium discoideum* (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Gloeckner G., Eichinger L., Szafanski K., Pachebat J., Dear P.,
 RA Lehmann R., Baugart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and Analysis of Chromosome 2 of *Dictyostelium*.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC115598; AAL92315.1;
 KW Hypothetical protein.
 SQ SEQUENCE 798 AA; 93093 MW; 69DC211E3E543965 CRC64;

Query Match 91.7%; Score 33; DB 5; Length 798;
 Best Local Similarity 83.3%; Pred. NO. 99;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 Db 501 FGPKNF 506

RESULT 7
 096316

ID 096316 PRELIMINARY; PRT; 1518 AA.
 AC 096316;
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE High voltage-activated calcium channel Cav2B (Fragment).
 OS *Schistosoma mansoni* (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigoidida;
 CC Schistosomatoidea; Schistosomatidae; Schistosoma.
 OX NCBI_TaxID=6183;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kohn A.B., Lea J.M., Roberts-Mistery J.M., Anderson P.A.V.,
 RA Greenberg R.M.;
 RT "Structure of Three High Voltage-Activated Calcium Channel Alpha-1
 RT Subunits from *Schistosoma mansoni*.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF361885; AAK84313.1;
 DR InterPro; IPR001682; Ca/Na.pore.
 DR InterPro; IPR002111; Cat_channel_TrpL.
 DR InterPro; IPR000636; M-channel_nlg.
 DR Pfam; PF00520; Ion_trans_4.
 DR PRINTS; PR01630; LYDCCALPHAL.
 FT NON_TER
 SQ SEQUENCE 1518 AA; 174745 MW; 43471939430A2C0C CRC64;

Query Match 91.7%; Score 33; DB 5; Length 1518;
 Best Local Similarity 83.3%; Pred. NO. 1.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 Db 1179 FGPKNF 1184

RESULT 8
 ID 097N51 PRELIMINARY; PRT; 279 AA.
 AC 097N51;
 DT 01-OCT-2001 (TRENBLREL. 18, Created)
 DT 01-OCT-2001 (TRENBLREL. 18, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE ABC transporter, ATP-binding protein.
 GN SP2220.
 OS *Streptococcus pneumoniae*.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 CC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TiGR4;
 RX MEDLINE-21357209; PubMed-11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., Deboy R.T., Hatt D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Redune D.,
 RA Holtzapple E., Khouri H., Wolf A.M., Uterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of *Streptococcus*
 RT *pneumoniae*.";
 RL Science 293:498-506(2001).
 DR EMBL; AE007509; AAK76268.1;
 DR TiGR; SP2220;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transportr.
 DR Pfam; PF00005; ABC_tran; 1.
 DR Prodom; PD000006; ABC_transportr; 1.
 DR SMART; SM00382; AAA_1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Complete proteome.
 SQ SEQUENCE 279 AA; 30525 MW; E051E473D082AD32 CRC64;

Query Match 88.9%; Score 32; DB 16; Length 279;
 Best Local Similarity 83.3%; Pred. No. 56;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 DB 109 FGPKNF 114

RESULT 9

099X12 PRELIMINARY; PRT; 280 AA.
 AC 099X12;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE Putative ABC transporter (ATP-binding protein).
 GN SPY2194.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
 RX MEDLINE=21192684; PubMed=11296296;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RT Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 (ABC TRANSPORTERS).
 DR EMBL: AE006636; AAK34822.1; -.
 DR HSSP: Q58663; 1G6H.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transporter.
 DR Pfam: PF00005; ABC_tran; 1.
 DR ProDom: PD000006; ABC_transport; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transport; Complete proteome.
 SQ SEQUENCE 280 AA; 30863 MW; DBE44F6F8780C7D5 CRC64;

Query Match 88.9%; Score 32; DB 16; Length 280;
 Best Local Similarity 83.3%; Pred. No. 56;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 DB 109 FGPKNF 114

RESULT 10

09C1S8 PRELIMINARY; PRT; 288 AA.
 AC 09C1S8;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE ABC transporter ATP-binding protein.
 GN YCHL OR IL0278.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21255186; PubMed=11337471;
 RA Bolotin A., Wincker P., Manger S., Jallion O., Malarne K.;

RA Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403.";
 RL Genome Res. 11:731-753(2001).
 DR EMBL: AE006265; AAK04376.1; -.
 DR HSSP: Q58663; 1G6J.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transporter.
 DR Pfam: PF00005; ABC_tran; 1.
 DR ProDom: PD000006; ABC_transport; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Complete proteome.
 SQ SEQUENCE 288 AA; 31848 MW; 9326C417DE7AEF22 CRC64;

Query Match 88.9%; Score 32; DB 16; Length 288;
 Best Local Similarity 83.3%; Pred. No. 57;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 DB 108 FGPKNF 113

RESULT 11

08R9L8 PRELIMINARY; PRT; 569 AA.
 AC 08R9L8;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DE ABC-type cobalt transport system, ATPase component.
 GN CB102 OR TRE1589.
 OS Thermobacterium tengcongensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermobacteriales; Thermobacteriaceae; Thermobacter.
 OX NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4T / JCM11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of T. tengcongensis genome.";
 RL Genome Res. 12:689-700(2002).
 DR EMBL: AE013114; AAM24793.1; -.
 KW Complete proteome.
 SQ SEQUENCE 569 AA; 64174 MW; FA59B5F9C632F2C1 CRC64;

Query Match 88.9%; Score 32; DB 16; Length 569;
 Best Local Similarity 83.3%; Pred. No. 1,2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 DB 113 FGPKNF 118

RESULT 12

09SUE7 PRELIMINARY; PRT; 912 AA.
 AC 09SUE7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 103.2 kDa protein.
 GN T13J8.20 OR ATG627910.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Pohl T., Weizenegger T., Hohelsel J., Mewes H.W.,
 RA Mayer K.F.X., Schueller C.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Pohl T., Weizenegger T., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: CONTAINS 1 SET DOMAIN.
 DR EMBL; AL035524; CAB36760.1; -;
 DR EMBL; AL161572; CAB79593.1; -;
 DR InterPro: IPR000313; PWMF_domain.
 DR InterPro: IPR001214; SET.
 DR InterPro: IPR01965; ZnF_PHD.
 DR Pfam; PF00855; PWMF; 1.
 DR Pfam; PF00856; SET; 1.
 DR SMART; SM00249; PHD; 1.
 DR SMART; SM00293; PWMF; 1.
 DR SMART; SM00317; SET; 1.
 DR PROSITE; PS0280; SET; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 912 AA; 103217 MW; BF6117E02E5297DE CRC64;

Query Match
 Best Local Similarity 88.9%; Score 32; DB 10; Length 912;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FGPKNF 6
 DB 200 FGPKNF 205

RESULT 13
 ID Q8W6C3 PRELIMINARY; PRT; 1206 AA.
 AC Q8W6C3;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE DNA polymerase III alpha subunit.
 GN Dpl.
 OS Saccharomonospora phage PIS 136.
 OC Viruses.
 OX NCBI_TaxID-182851;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Khandrika L., Soni V., Agrawal P.;
 RT "Putative DNA polymerase III alpha subunit of bacteriophage PIS 136
 submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF559434; AAL66178.1; -;
 DR InterPro: IPR004013; PHP_C.
 DR InterPro: IPR003141; PHP_N.
 DR InterPro: IPR004805; PolC_alpha.
 DR InterPro: IPR004365; trna_anti.
 DR Pfam; PF02811; PHP_C; 1.
 DR Pfam; PF02231; PHP_N; 1.
 DR Pfam; PF01336; trna_anti; 1.
 DR SMART; SM00481; POLITAC; 1.
 DR TIGRfams; TIGR00594; POLC; 1.
 SQ SEQUENCE 1206 AA; 132633 MW; 64E580CF91C63BBB CRC64;

Query Match 88.9%; Score 32; DB 9; Length 1206;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FGPKNF 6
 DB 251 FGPKNF 256

RESULT 14
 ID Q9ZUP7 PRELIMINARY; PRT; 307 AA.
 AC Q9ZUP7;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE En/Spm transposon protein (mosaic-like protein).
 GN At2G12300.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID-3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RX MEDLINE-20083487; PubMed-10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldlyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.-J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC005897; AAC97242.1; -;
 DR InterPro: IPR004252; Transposase_24.
 DR Pfam; PF03004; Transposase_24; 1.
 SQ SEQUENCE 307 AA; 35032 MW; 12B7D900DC3144C1 CRC64;

Query Match 86.1%; Score 31; DB 10; Length 307;
 Best Local Similarity 83.3%; Pred. No. 1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FGPKNF 6
 DB 150 FGPKNF 155

RESULT 15
 ID Q9C6S0 PRELIMINARY; PRT; 593 AA.
 AC Q9C6S0;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Cytochrome P450, putative.
 GN F5M6.19.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID-3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RX MEDLINE-21016719; PubMed-11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,

RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gall J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetskaya I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luos J.S., Maiti R., Marziani A.,
 RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utechtack T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana."
 RL Nature 408:816-820(2000).
 CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL: AC079041; AAG50718.1; -.
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
 KW Heme: Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 593 AA; 66643 MW; E80CBE9B8B2BD199 CRC64;

Query Match 86.1%; Score 31; DB 10; Length 593;
 Best Local Similarity 83.3%; Pred. No. 2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FGPKNF 6
 Db 145 FGPKSF 150

Search completed: July 18, 2003, 16:03:45
 Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 16:00:24 : Search time 19.3684 Seconds
(without alignments)
12.153 Million cell updates/sec

Title: US-10-054-968-8

Perfect score: 35

Sequence: 1 RVALAGVL 8

Scoring table: BLOSUM62

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued_Patents_AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	290	US-09-134-001C-3273	Sequence 3273, Ap
2	28	80.0	214	US-08-217-327-4	Sequence 4, Appli
3	28	80.0	491	US-09-134-001C-4612	Sequence 4612, Ap
4	28	80.0	493	US-09-177-349-5	Sequence 5, Appli
5	28	80.0	640	US-09-177-349-4	Sequence 4, Appli
6	28	80.0	956	US-08-185-232A-2	Sequence 2, Appli
7	28	80.0	956	US-08-416-523-2	Sequence 2, Appli
8	28	80.0	956	US-08-789-478-2	Sequence 2, Appli
9	27	77.1	272	PCT-US95-02455-4	Sequence 4, Appli
10	27	77.1	286	US-08-964-127-4	Sequence 4, Appli
11	27	77.1	286	US-09-486-692-4	Sequence 4, Appli
12	27	77.1	310	US-09-794-336-4	Sequence 4, Appli
13	27	77.1	423	US-09-134-001C-3599	Sequence 3599, Ap
14	27	77.1	479	US-08-583-318-5	Sequence 5, Appli
15	27	77.1	520	US-08-964-127-2	Sequence 2, Appli
16	27	77.1	520	US-09-496-692-2	Sequence 2, Appli
17	27	77.1	765	US-08-737-109-11	Sequence 11, Appli
18	27	77.1	853	US-08-689-103B-10	Sequence 10, Appli
19	27	77.1	853	US-09-229-059-10	Sequence 10, Appli
20	27	77.1	908	US-08-699-103B-12	Sequence 12, Appli
21	27	77.1	908	US-09-229-059-12	Sequence 12, Appli
22	27	77.1	2254	US-08-677-010-3	Sequence 3, Appli
23	27	77.1	2254	US-08-790-519-3	Sequence 3, Appli
24	27	77.1	3170	US-09-036-987A-4	Sequence 4, Appli
25	27	77.1	3170	US-09-370-700-4	Sequence 4, Appli
26	26	74.3	24	US-08-103-445-18	Sequence 18, Appli
27	26	74.3	24	US-08-461-690B-18	Sequence 18, Appli

28	26	74.3	68	US-09-243-675-2	Sequence 2, Appli
29	26	74.3	68	US-09-243-675-3	Sequence 3, Appli
30	26	74.3	75	US-09-243-675-1	Sequence 1, Appli
31	26	74.3	199	US-08-831-575-9	Sequence 9, Appli
32	26	74.3	260	US-09-796-774-1	Sequence 1, Appli
33	26	74.3	260	US-08-663-831-1	Sequence 1, Appli
34	26	74.3	260	US-08-663-831-2	Sequence 2, Appli
35	26	74.3	260	US-08-663-831-3	Sequence 3, Appli
36	26	74.3	260	US-08-663-831-4	Sequence 4, Appli
37	26	74.3	260	US-08-663-831-5	Sequence 5, Appli
38	26	74.3	260	US-08-663-831-6	Sequence 6, Appli
39	26	74.3	260	US-08-663-831-7	Sequence 7, Appli
40	26	74.3	260	US-08-663-831-8	Sequence 8, Appli
41	26	74.3	260	US-08-663-831-9	Sequence 9, Appli
42	26	74.3	260	US-08-663-831-10	Sequence 10, Appli
43	26	74.3	260	US-08-663-831-11	Sequence 11, Appli
44	26	74.3	260	US-08-663-831-12	Sequence 12, Appli
45	26	74.3	260	US-08-663-831-13	Sequence 13, Appli

ALIGNMENTS

```
RESULT 1
US-09-134-001C-3273
; Sequence 3273, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: GNC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR APPLICATION NUMBER: 1998-08-13
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3273
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3273

Query Match      100.0%; Score 35; DB 4; Length 290;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RVALAGVL 8
Db      171 RVALAGVL 178

RESULT 2
US-08-217-327-4
; Sequence 4, Application US/08217327
; Patent No. 5474925
; GENERAL INFORMATION:
; APPLICANT: John, Malyakal E
; APPLICANT: Barton, Kenneth A
; TITLE OF INVENTION: Immobilized Proteins in Cotton Fiber
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: Quarles and Brady
; STREET: P.O. Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/217,327
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/812,233
FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 1122990831
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-217-327-4

Query Match 80.0%; Score 28; DB 1; Length 214;
Best Local Similarity 71.4%; Pred. No. 77;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 VALAGVL 8
DB 16 IAVAGVL 22

RESULT 3
US-09-134-001C-4612
Sequence 4612, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4612
LENGTH: 491
TYPE: PRN
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4612

Query Match 80.0%; Score 28; DB 4; Length 491;
Best Local Similarity 85.7%; Pred. No. 1,9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VALAGVL 8
DB 436 VALAGVL 442

RESULT 4
US-09-177-349-5
Sequence 5, Application US/09177349
Patent No. 6268201
GENERAL INFORMATION:
APPLICANT: Alland, David
APPLICANT: Bloom, Barry R.
APPLICANT: Jacobs Jr., William R.
TITLE OF INVENTION: IN1B, IN1A AND IN1C GENES OF MYCOBACTERIA AND METHODS

TITLE OF INVENTION: OF USE
FILE REFERENCE: 96700/491
CURRENT APPLICATION NUMBER: US/09/177,349
CURRENT FILING DATE: 1998-10-23
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 493
TYPE: PRN
ORGANISM: Mycobacterium tuberculosis
US-09-177-349-5

Query Match 80.0%; Score 28; DB 4; Length 493;
Best Local Similarity 62.5%; Pred. No. 1,9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVALAGVL 8
DB 46 RIALAGVL 53

RESULT 5
US-09-177-349-4
Sequence 4, Application US/09177349
Patent No. 6268201
GENERAL INFORMATION:
APPLICANT: Alland, David
APPLICANT: Bloom, Barry R.
APPLICANT: Jacobs Jr., William R.
TITLE OF INVENTION: IN1B, IN1A AND IN1C GENES OF MYCOBACTERIA AND METHODS
TITLE OF INVENTION: OF USE
FILE REFERENCE: 96700/491
CURRENT APPLICATION NUMBER: US/09/177,349
CURRENT FILING DATE: 1998-10-23
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 640
TYPE: PRN
ORGANISM: Mycobacterium tuberculosis
US-09-177-349-4

Query Match 80.0%; Score 28; DB 4; Length 640;
Best Local Similarity 75.0%; Pred. No. 2,6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVALAGVL 8
DB 72 RVALAGVL 79

RESULT 6
US-08-185-232A-2
Sequence 2, Application US/08185232A
Patent No. 5576205
GENERAL INFORMATION:
APPLICANT: KAMBOJ, Rajender
APPLICANT: NUTT, Stephen
APPLICANT: SHEKTER, Lee
APPLICANT: WOSNICK, Michael
TITLE OF INVENTION: KAINATE-BINDING HUMAN CNS RECEPTORS OF
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K. Street, N.W. Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,232A
FILING DATE: 24-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/750,090
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/150 ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 956 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-185-232A-2

Query Match 80.0%; Score 28; DB 1; Length 956;
Best Local Similarity 71.4%; Pred. No. 4.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 VA1AGVL 8
11:11:1
Db 148 VAVAGIL 154

RESULT 7
US-08-416-523-2
Sequence 2, Application US/08416523
Patent No. 5616481
GENERAL INFORMATION:
APPLICANT: KAMBOJ, Rajender
APPLICANT: NUTT, Stephen
APPLICANT: SHEKTER, Lee
APPLICANT: MOSNICK, Michael
TITLE OF INVENTION: KAINATE-BINDING HUMAN CNS RECEPTORS OF
TITLE OF INVENTION: THE EAAL FAMILY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,523
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,441
FILING DATE: 15-JUL-1993
APPLICATION NUMBER: US 07/750,090
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/219 ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 956 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-416-523-2

Query Match 80.0%; Score 28; DB 1; Length 956;
Best Local Similarity 71.4%; Pred. No. 4.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 VA1AGVL 8
11:11:1
Db 148 VAVAGIL 154

RESULT 8
US-08-789-478-2
Sequence 2, Application US/08789478
Patent No. 6013768
GENERAL INFORMATION:
APPLICANT: KAMBOJ, Rajender
APPLICANT: NUTT, Stephen
APPLICANT: SHEKTER, Lee
APPLICANT: MOSNICK, Michael
TITLE OF INVENTION: KAINATE-BINDING HUMAN CNS RECEPTORS OF
TITLE OF INVENTION: THE EAAL FAMILY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/789,478
FILING DATE: 27-JAN-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,441
FILING DATE: 15-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/750,090
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/303 ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 956 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-789-478-2

Query Match 80.0%; Score 28; DB 3; Length 956;
Best Local Similarity 71.4%; Pred. No. 4.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 VA1AGVL 8
11:11:1
Db 148 VAVAGIL 154

RESULT 9
PCT-US95-02455-4
; Sequence 4, Application PC/TUS9502455
; GENERAL INFORMATION:
; APPLICANT: JACOBS et al
; TITLE OF INVENTION: GENE FOR MYCOBACTERIAL
; TITLE OF INVENTION: DIAMINOPIMELIC ACID
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amster, Rochstein & Ebnstein
; STREET: 90 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 MB
; MEDIUM TYPE: storage diskette
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02455
; FILING DATE: Not Yet Assigned
; CLASSIFICATION:
; PRIOR APPLICATION DATA: none
; APPLICATION NUMBER: none
; FILING DATE: none
; ATTORNEY/AGENT INFORMATION:
; NAME: George, Kenneth P.
; REGISTRATION NUMBER: 30,259
; REFERENCE/DOCKET NUMBER: 96700/358
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: dnpb protein
; HYPOTHEICAL: no
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE: E coli
; ORGANISM: E coli
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Bouvier et al
; TITLE:
; JOURNAL: J Biol Chem
; VOLUME: 259
; PAGES: 14829-14834

DATE: (1984)
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
PCT-US95-02455-4
Query Match 77.1%; Score 27; DB 5; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RVAIAG 6
DB 7 RVAIAG 12
RESULT 10
US-08-964-127-4
; Sequence 4, Application US/08964127
; Patent No. 6277565
; GENERAL INFORMATION:
; APPLICANT: Grandearl, Andrew David John
; TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,127
; FILING DATE: 06-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crews, Ph.D., L. Lee
; REGISTRATION NUMBER: P-43,567
; REFERENCE/DOCKET NUMBER: 07334/038001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-8906
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-964-127-4
Query Match 77.1%; Score 27; DB 4; Length 286;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 RVAIAGVL 8
DB 187 RVAIAGEL 194
RESULT 11
US-09-496-692-4
; Sequence 4, Application US/09496692
; Patent No. 6313271

GENERAL INFORMATION:
APPLICANT: Grandearl, Andrew David John
TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/496,692
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/964,127
FILING DATE: 06-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Crews, Ph.D., L. Lee
REGISTRATION NUMBER: P-43,567
REFERENCE/DOCKET NUMBER: 07334/038001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 286 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-496-692-4

Query Match 77.1%; Score 27; DB 4; Length 286;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RVATAGYL 8
Db 187 RVALAGEL 194

RESULT 12
US-09-794-236-4
Sequence 4, Application US/09794236
Patent No. 6337069
GENERAL INFORMATION:
APPLICANT: Grouzmann, Eric
APPLICANT: Lacroix, Jean-Silvain
APPLICANT: Morod, Michel
TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
FILE REFERENCE: 81985/276823
CURRENT APPLICATION NUMBER: US/09/794,236
CURRENT FILING DATE: 2001-02-28
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 310
TYPE: PRT
ORGANISM: Homo sapiens
US-09-794-236-4

Query Match 77.1%; Score 27; DB 4; Length 310;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RVATAG 6
Db 185 RVALAG 190

RESULT 13
US-09-134-001C-3599
Sequence 3599, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NOCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3599
LENGTH: 423
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3599

Query Match 77.1%; Score 27; DB 4; Length 423;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RVATAGYL 8
Db 247 RSAVAGIL 254

RESULT 14
US-08-583-318-5
Sequence 5, Application US/08583318
Patent No. 5693483
GENERAL INFORMATION:
APPLICANT: Staunton, Donald
APPLICANT: Harris, Edith
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
TITLE OF INVENTION: Binding
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/583,318
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/33050
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 479 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-583-318-5

Query Match 77.1%; Score 27; DB 1; Length 479;
Best Local Similarity 62.5%; Pred. No. 3.1e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVALAGVL 8
:|:|:|:|
Db 361 QVAVAGLL 368

RESULT 15

US-08-964-127-2
; Sequence 2, Application US/08964127
; Patent No. 6277565
; GENERAL INFORMATION:
; APPLICANT: Grandearl, Andrew David John
; TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,127
; FILING DATE: 06-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crews, Ph.D., L. Lee
; REGISTRATION NUMBER: P-43,567
; REFERENCE/DOCKET NUMBER: 07334/038001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-964-127-2

Query Match 77.1%; Score 27; DB 4; Length 520;
Best Local Similarity 75.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVALAGVL 8
:|:|:|:|
Db 187 RVALAGEL 194

Search completed: July 18, 2003, 16:05:19
Job time : 20.3684 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 18, 2003, 15:56:19 ; Search time 23.6842 Seconds
(without alignments)
28.131 Million cell updates/sec

Title: US-10-054-968-6
Perfect score: 23
Sequence: 1 KVGIV 5

Scoring table: BIOSM62
Gap 10.0, Gape 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_101002.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	100.0	5	AA44379	Vancomycin resista
2	23	100.0	62	AB803510	Human musculoskele
3	23	100.0	67	AA621853	Arabidopsis thalia
4	23	100.0	68	ABP00046	Human ORFX protein
5	23	100.0	76	AAW20271	H. pylori surface
6	23	100.0	76	AA615623	Arabidopsis thalia
7	23	100.0	79	AA626427	Arabidopsis thalia
8	23	100.0	79	ABG09451	Novel human diagno
9	23	100.0	85	AA616008	Arabidopsis thalia
10	23	100.0	90	ABG18024	Novel human diagno

11	23	100.0	92	AA621224	Arabidopsis thalia
12	23	100.0	92	AA634700	Arabidopsis thalia
13	23	100.0	92	AB861844	Drosophila melanog
14	23	100.0	95	AA857210	Human prostate can
15	23	100.0	101	AA002049	Human polypeptide
16	23	100.0	108	ABG09452	Novel human diagno
17	23	100.0	108	ABG60148	Human DTRP polype
18	23	100.0	111	AA615858	Arabidopsis thalia
19	23	100.0	111	ABP42814	Human ovarian anti
20	23	100.0	111	ABP42868	Human ovarian anti
21	23	100.0	113	AA857211	Human prostate can
22	23	100.0	116	AA600228	E faecalis EF118 p
23	23	100.0	116	ABP43447	Human ovarian anti
24	23	100.0	120	ABP41368	Human ovarian anti
25	23	100.0	127	AA818071	Pinus radiata furo
26	23	100.0	127	ABP03341	Human ORFX protein
27	23	100.0	128	AA857209	Human prostate can
28	23	100.0	140	ABP07556	Human ORFX protein
29	23	100.0	142	ABG22909	Novel human diagno
30	23	100.0	143	AA617980	Arabidopsis thalia
31	23	100.0	166	AA628351	Arabidopsis thalia
32	23	100.0	172	AA651027	A. gossypii rib5 p
33	23	100.0	172	AA651031	A. gossypii rib4 p
34	23	100.0	176	AA610968	H. pylori ORF hp5p
35	23	100.0	180	AA644388	Vancomycin resista
36	23	100.0	184	AAW20635	H. pylori transmem
37	23	100.0	184	AAU23460	Novel human enzyme
38	23	100.0	186	AA624068	Arabidopsis thalia
39	23	100.0	203	AA624067	Arabidopsis thalia
40	23	100.0	213	AAW20492	H. pylori cell env
41	23	100.0	216	AAW20711	H. pylori transpor
42	23	100.0	224	AA628223	Multi-drug resista
43	23	100.0	224	AA628225	Multi-drug resista
44	23	100.0	224	AA628227	Multi-drug resista
45	23	100.0	229	AB847434	Listeria monocytog

ALIGNMENTS

RESULT 1
AAW44379 standard; peptide: 5 AA.
ID AAW44379
XX
AC AAW44379;
XX
DT 10-JUN-1998 (first entry)
XX
DE Vancomycin resistant Enterococcus faecium specific epitope SEQ ID NO:6.
XX
KW Vancomycin resistant Enterococcus faecium; VRE; epitope; immunogen;
KW methicillin resistant Staphylococcus aureus; MRSA; transporter protein.
XX
OS Enterococcus faecium.
XX
PN WO9801154-A2.
XX
PD 15-JAN-1998.
XX
PF 07-JUL-1997; 97WO-GB01830.
XX
PR 06-JUL-1996; 96GB-0014274.
XX
PA (NEUT-) NEUTEC PHARM PLC.
XX
PI Burnie JP, Matthews RC;
XX
DR WPI; 1998-100821/09.
XX
PT Treating and diagnosing bacterial and fungal infection with ABC
PT transporter protein - or neutralising or binding agents, and new
PT Staphylococcal proteins, particularly for infections caused by drug
PT resistant Staphylococci and Enterococci

XX Claim 9; Page 38; 62pp; English.
PS
XX
CC The present sequence represents a vancomycin resistant Enterococcus
CC faecium (VRE) specific epitope. The present invention describes the
CC therapeutic or diagnostic use of a bacterial or fungal ABC transporter
CC protein, or its immunogenic fragments, in human or veterinary medicine.
CC Binding agents, specific for an ABC transporter protein, are used in
CC standard binding assays to detect the ABC transporter protein, while
CC ABC transporter proteins are used similarly to detect specific
CC antibodies in patient sera, specifically for diagnosis of Staphylococci
CC and Enterococci infections. Neutralising agents specific for ABC
CC transporter proteins, specifically antibodies (Ab), are used to treat
CC such infections (optionally when coupled to a drug) while ABC
CC transporter proteins are used as immunogens to protect against
CC infection. Ab raised against ABC transporter proteins are effective
CC against multiple drug resistant strains of bacteria (specifically
CC vancomycin and/or methicillin resistant strains) for which no drug
CC therapy is available.
XX
SQ Sequence 5 AA;
Query Match 100.0%; Score 23; DB 19; Length 5;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVGIV 5
Db 1 KVGIV 5
RESULT 2
ABBO3510 ABB03510 standard; Protein; 62 AA.
XX
AC ABB03510;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polypeptide SEQ ID NO 1457.
XX
KM Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KM antiasthmatic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KM vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KM neurological disease; infection; human; secreted protein;
KM musculoskeletal system.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239360.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0246475.

CC	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249267.
PR	17-NOV-2000;	2000US-0249287.
PR	17-NOV-2000;	2000US-0249289.
PR	01-DEC-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251899.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX	WPI: 2001-451937/48.	
DR	N-PSDB; AAL35092.	
XX		
PT	Isolated polypeptide for treating, preventing and/or prognosing	
PT	disorders related to the musculoskeletal system including	
PT	musculoskeletal cancers and also for testing and detection e.g.	
PT	diagnosis -	
XX		
PS	Claim 11; SEQ ID NO 1457; 781pp + Sequence Listing; English.	
XX		
CC	The invention relates to novel genes (AAL34669-AAL37666) and proteins	
CC	(AAB03087-AAB04109) associated with the musculoskeletal system useful	
CC	for preventing, treating or ameliorating medical conditions e.g. by	
CC	protein or gene therapy. The genes are isolated from a range of human	
CC	tissues disclosed in the specification. The nucleic acids, proteins,	
CC	antibodies and (ant)agonists are useful in the diagnosis, treatment	
CC	and prevention of: (a) cancer, e.g. breast and ovarian cancer and	
CC	other cancers of the adrenal gland, bone, bone marrow, breast,	
CC	gastrointestinal tract, liver, lung, or urogenital; (b) immune	
CC	disorders e.g. Addison's disease, allergies, autoimmune haemolytic	
CC	anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,	
CC	multiple sclerosis, rheumatoid arthritis and ulcerative colitis;	

Query Match	Best Local Similarity	Score 23;	DB 22;	Length 62;
Matches	5;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 KVGIV 5			
Db	30 KVGIV 34			
<p>RESULT 3</p> <p>AAAG21853</p> <p>ID AAAG21853 standard; Protein; 67 AA.</p> <p>AC AAAG21853;</p> <p>AD</p> <p>AE 17-OCT-2000 (first entry)</p> <p>AF</p> <p>AG Arabidopsis thaliana protein fragment SEQ ID NO: 24561.</p> <p>AH</p> <p>AI Protein identification; signal transduction pathway; metabolic pathway; hydridisation assay; genetic mapping; gene expression control; promoter; termination sequence.</p> <p>AK</p> <p>AL Arabidopsis thaliana.</p> <p>AM</p> <p>AN EP1033405-A2.</p> <p>AO</p> <p>AP</p> <p>AQ</p> <p>AR</p> <p>AS</p> <p>AT</p> <p>AV</p> <p>AW</p> <p>AX</p> <p>AY</p> <p>AZ</p> <p>BA</p> <p>BB</p> <p>BC</p> <p>BD</p> <p>BE</p> <p>BF</p> <p>BG</p> <p>BH</p> <p>BI</p> <p>BJ</p> <p>BK</p> <p>BL</p> <p>BM</p> <p>BN</p> <p>BO</p> <p>BP</p> <p>BQ</p> <p>BR</p> <p>BS</p> <p>BT</p> <p>BU</p> <p>BV</p> <p>BW</p> <p>BX</p> <p>BY</p> <p>BZ</p> <p>CA</p> <p>CB</p> <p>CC</p> <p>CD</p> <p>CE</p> <p>CF</p> <p>CG</p> <p>CH</p> <p>CI</p> <p>CJ</p> <p>CK</p> <p>CL</p> <p>CM</p> <p>CN</p> <p>CO</p> <p>CP</p> <p>CQ</p> <p>CR</p> <p>CS</p> <p>CT</p> <p>CU</p> <p>CV</p> <p>CW</p> <p>CX</p> <p>CY</p> <p>CZ</p> <p>DA</p> <p>DB</p> <p>DC</p> <p>DD</p> <p>DE</p> <p>DF</p> <p>DG</p> <p>DH</p> <p>DI</p> <p>DJ</p> <p>DK</p> <p>DL</p> <p>DM</p> <p>DN</p> <p>DO</p> <p>DP</p> <p>DQ</p> <p>DR</p> <p>DS</p> <p>DT</p> <p>DU</p> <p>DV</p> <p>DW</p> <p>DX</p> <p>DY</p> <p>DZ</p> <p>EA</p> <p>EB</p> <p>EC</p> <p>ED</p> <p>EE</p> <p>EF</p> <p>EG</p> <p>EH</p> <p>EI</p> <p>EJ</p> <p>EK</p> <p>EL</p> <p>EM</p> <p>EN</p> <p>EO</p> <p>EP</p> <p>EQ</p> <p>ER</p> <p>ES</p> <p>ET</p> <p>EU</p> <p>EV</p> <p>EW</p> <p>EX</p> <p>EY</p> <p>EZ</p> <p>FA</p> <p>FB</p> <p>FC</p> <p>FD</p> <p>FE</p> <p>FF</p> <p>FG</p> <p>FH</p> <p>FI</p> <p>FJ</p> <p>FK</p> <p>FL</p> <p>FM</p> <p>FN</p> <p>FO</p> <p>FP</p> <p>FQ</p> <p>FR</p> <p>FS</p> <p>FT</p> <p>FU</p> <p>FV</p> <p>FW</p> <p>FX</p> <p>FY</p> <p>FZ</p> <p>GA</p> <p>GB</p> <p>GC</p> <p>GD</p> <p>GE</p> <p>GF</p> <p>GG</p> <p>GH</p> <p>GI</p> <p>GJ</p> <p>GK</p> <p>GL</p> <p>GM</p> <p>GN</p> <p>GO</p> <p>GP</p> <p>GQ</p> <p>GR</p> <p>GS</p> <p>GT</p> <p>GU</p> <p>GV</p> <p>GW</p> <p>GX</p> <p>GY</p> <p>GZ</p> <p>HA</p> <p>HB</p> <p>HC</p> <p>HD</p> <p>HE</p> <p>HF</p> <p>HG</p> <p>HH</p> <p>HI</p> <p>HJ</p> <p>HK</p> <p>HL</p> <p>HM</p> <p>HN</p> <p>HO</p> <p>HP</p> <p>HQ</p> <p>HR</p> <p>HS</p> <p>HT</p> <p>HU</p> <p>HV</p> <p>HW</p> <p>HX</p> <p>HY</p> <p>HZ</p> <p>IA</p> <p>IB</p> <p>IC</p> <p>ID</p> <p>IE</p> <p>IF</p> <p>IG</p> <p>IH</p> <p>II</p> <p>IJ</p> <p>IK</p> <p>IL</p> <p>IM</p> <p>IN</p> <p>IO</p> <p>IP</p> <p>IQ</p> <p>IR</p> <p>IS</p> <p>IT</p> <p>IU</p> <p>IV</p> <p>IW</p> <p>IX</p> <p>IY</p> <p>IZ</p> <p>JA</p> <p>JB</p> <p>JC</p> <p>JD</p> <p>JE</p> <p>JF</p> <p>JG</p> <p>JH</p> <p>JI</p> <p>BJ</p> <p>JK</p> <p>JL</p> <p>JM</p> <p>JN</p> <p>JO</p> <p>JP</p> <p>JQ</p> <p>JR</p> <p>JS</p> <p>JT</p> <p>JU</p> <p>JV</p> <p>JW</p> <p>JX</p> <p>JY</p> <p>JZ</p> <p>KA</p> <p>KB</p> <p>KC</p> <p>KD</p> <p>KE</p> <p>KF</p> <p>KG</p> <p>KH</p> <p>KI</p> <p>KJ</p> <p>KL</p> <p>KM</p> <p>KN</p> <p>KO</p> <p>KP</p> <p>KQ</p> <p>KR</p> <p>KS</p> <p>KT</p> <p>KU</p> <p>KV</p> <p>KW</p> <p>KX</p> <p>KY</p> <p>KZ</p> <p>LA</p> <p>LB</p> <p>LC</p> <p>LD</p> <p>LE</p> <p>LF</p> <p>LG</p> <p>LH</p> <p>LI</p> <p>LJ</p> <p>LK</p> <p>LL</p> <p>LM</p> <p>LN</p> <p>LO</p> <p>LP</p> <p>LQ</p> <p>LR</p> <p>LS</p> <p>LT</p> <p>LU</p> <p>LV</p> <p>LW</p> <p>LX</p> <p>LY</p> <p>LZ</p> <p>MA</p> <p>MB</p> <p>MC</p> <p>MD</p> <p>ME</p> <p>MF</p> <p>MG</p> <p>MH</p> <p>MI</p> <p>MJ</p> <p>MK</p> <p>ML</p> <p>MM</p> <p>MN</p> <p>MO</p> <p>MP</p> <p>MQ</p> <p>MR</p> <p>MS</p> <p>MT</p> <p>MU</p> <p>MV</p> <p>MW</p> <p>MX</p> <p>MY</p> <p>MZ</p> <p>NA</p> <p>NB</p> <p>NC</p> <p>ND</p> <p>NE</p> <p>NF</p> <p>NG</p> <p>NH</p> <p>NI</p> <p>NJ</p> <p>NK</p> <p>NL</p> <p>NM</p> <p>NO</p> <p>NP</p> <p>NQ</p> <p>NR</p> <p>NS</p> <p>NT</p> <p>NU</p> <p>NV</p> <p>NW</p> <p>NX</p> <p>NY</p> <p>NZ</p> <p>OA</p> <p>OB</p> <p>OC</p> <p>OD</p> <p>OE</p> <p>OF</p> <p>OG</p> <p>OH</p> <p>OI</p> <p>OJ</p> <p>OK</p> <p>OL</p> <p>OM</p> <p>ON</p> <p>OO</p> <p>OP</p> <p>OQ</p> <p>OR</p> <p>OS</p> <p>OT</p> <p>OU</p> <p>OV</p> <p>OW</p> <p>OX</p> <p>OY</p> <p>OZ</p> <p>PA</p> <p>PB</p> <p>PC</p> <p>PD</p> <p>PE</p> <p>PF</p> <p>PG</p> <p>PH</p> <p>PI</p> <p>PJ</p> <p>PK</p> <p>PL</p> <p>PM</p> <p>PN</p> <p>PO</p> <p>PP</p> <p>PQ</p> <p>PR</p> <p>PS</p> <p>PT</p> <p>PV</p> <p>PW</p> <p>PX</p> <p>PY</p> <p>PZ</p> <p>QA</p> <p>QB</p> <p>QC</p> <p>QD</p> <p>QE</p> <p>QF</p> <p>QG</p> <p>QH</p> <p>QI</p> <p>QJ</p> <p>QK</p> <p>QL</p> <p>QM</p> <p>QN</p> <p>QO</p> <p>QP</p> <p>QQ</p> <p>QR</p> <p>QS</p> <p>QT</p> <p>QU</p> <p>QV</p> <p>QW</p> <p>QX</p> <p>QY</p> <p>QZ</p> <p>RA</p> <p>RB</p> <p>RC</p> <p>RD</p> <p>RE</p> <p>RF</p> <p>RG</p> <p>RH</p> <p>RI</p> <p>RJ</p> <p>RK</p> <p>RL</p> <p>RM</p> <p>RO</p> <p>RP</p> <p>RQ</p> <p>RR</p> <p>RS</p> <p>RT</p> <p>RU</p> <p>RV</p> <p>RW</p> <p>RX</p> <p>RY</p> <p>RZ</p> <p>SA</p> <p>SB</p> <p>SC</p> <p>SD</p> <p>SE</p> <p>SF</p> <p>SG</p> <p>SH</p> <p>SI</p> <p>SJ</p> <p>SK</p> <p>SL</p> <p>SM</p> <p>SN</p> <p>SO</p> <p>SP</p> <p>SQ</p> <p>SR</p> <p>SS</p> <p>ST</p> <p>SU</p> <p>SV</p> <p>SW</p> <p>SX</p> <p>SY</p> <p>SZ</p> <p>TA</p> <p>TB</p> <p>TC</p> <p>TD</p> <p>TE</p> <p>TF</p> <p>TG</p> <p>TH</p> <p>TI</p> <p>TJ</p> <p>TK</p> <p>TL</p> <p>TM</p> <p>TN</p> <p>TO</p> <p>TP</p> <p>TQ</p> <p>TR</p> <p>TS</p> <p>TT</p> <p>TU</p> <p>TV</p> <p>TW</p> <p>TX</p> <p>TY</p> <p>TZ</p> <p>UA</p> <p>UB</p> <p>UC</p> <p>UD</p> <p>UE</p> <p>UF</p> <p>UG</p> <p>UH</p> <p>UI</p> <p>UJ</p> <p>UK</p> <p>UL</p> <p>UM</p> <p>UN</p> <p>UO</p> <p>UP</p> <p>UQ</p> <p>UR</p> <p>US</p> <p>UT</p> <p>UU</p> <p>UV</</p>				

PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match

Best Local Similarity 100.0%; Score 23; DB 21; Length 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
| | | | |
Db 7 KVGIV 11

RESULT 4
ABP00046
ID ABP00046 standard; protein; 68 AA.

AC ABP00046;

DT 24-JUN-2002 (first entry)

DE Human ORFX protein sequence SEQ ID NO:74.

KM Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KM hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KM degenerative disorder; osteoarthritis; neurodegenerative disorder;
KM cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KM hypertension; hypothyroidism; cholesterol ester storage disease;
KM immune deficiency; immune disorder; infectious disease;
KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KM myasthenia gravis.

OS Homo sapiens.

PN WO200192523-A2.

PD 06-DEC-2001.

PF 29-MAY-2001; 2001WO-US10836.

PR 30-MAY-2000; 2000US-206132P.

PR 29-AUG-2000; 2000US-228716P.

PA (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach MD;

DR WPI; 2002-106308/14.

DR N-PSDB; ABN15798.

Novel human polypeptides and polynucleotides useful for diagnosing,
preventing and treating cardiovascular disease, neurodegenerative,
hyperproliferative disorders and autoimmune disorders

Disclosure; SEQ ID 74; 1037pp; English.

The present invention describes substantially purified human proteins
(referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
in the specification). ABN15762 to ABN27252 encode the human ORFX
proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
treating or preventing a pathology associated with an ORFX-associated
disorder in humans, and in the manufacture of a medicament for treating a
syndrome associated with ORFX-associated disorder. ORFX polynucleotide
sequences can be used in gene therapy. ORFX sequences can be used in the
treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
osteoarthritis, neurodegenerative disorders, disorders related to organ
transplantation, cardiovascular diseases, diabetes mellitus, systemic
lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
storage disease, various immune deficiencies and disorders, infectious
diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
disease and autoimmune inflammatory eye disease. ORFX proteins are also
useful for treating burns, incisions, ulcers, for treating osteoporosis,

CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

SO Sequence 68 AA;

Query Match 100.0%; Score 23; DB 23; Length 68;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
| | | | |
Db 35 KVGIV 39

RESULT 5
AAW20271
ID AAW20271 standard; protein; 76 AA.

AC AAW20271;

DT 09-JUL-1997 (first entry)

DE H. pylori surface or membrane protein, 23867207.aa.

KM Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KM binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KM duodenal ulcer disease; chronic gastritis; diagnosis; envelope.

OS Helicobacter pylori.

PN WO9640893-A1.

PD 19-DEC-1996.

PF 06-JUN-1996; 96WO-US09122.

PR 01-APR-1996; 96US-0630405.

PR 07-JUN-1995; 95US-0487032.

PA (ASTRA) ASTRA AB.

PI Berglindh OT, Smith D, Mellgaard BL;

DR WPI; 1997-052306/05.

Novel Helicobacter pylori nucleic acid sequences and related
polypeptide(s) - useful for vaccines to treat or prevent H. pylori
infection, and to detect Helicobacter

Claim 73; Page 474; 1481pp; English.

The present sequence is a Helicobacter pylori surface or membrane protein
likely to contain at least two membrane spanning regions.
The protein may be used in a vaccine to prevent or treat H. pylori
infection or to identify H. pylori polypeptide binding compounds. To
useful as potential H. pylori life cycle activators or inhibitors.
The genomic sequence of H. pylori (ATCC 55679) was determined from
overlapping contigs generated by mechanically shearing the bacterial
DNA. The sequences were analysed for ORF of at least 180 nucleotides,
and the predicted coding regions defined by computer evaluation. To
identify likely H. pylori antigens for vaccine development, the amino
acid sequences predicted from various ORFs were analysed for significant
homology to other known or exported membrane proteins. Having identified
and determined the sequences of interest, particular regions can be
isolated from H. pylori by PCR amplification for recombinant polypeptide
production, e.g. in E. coli hosts.

Sequence 76 AA;
Query Match 100.0%; Score 23; DB 18; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
|||||
DB 6 KVGIV 10

RESULT 6
AAG15623
ID AAG15623 standard; Protein; 76 AA.
XX
AC AAG15623;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 15948.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128645.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 23; DB 21; Length 76;
Best Local Similarity 100.0%; Pred. NO. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KGVIV 5

Db 7 KGVIV 11
RESULT 7
AAG26427
ID AAG26427 standard; Protein: 79 AA.
XX
AC AAG26427;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 30879.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydrolisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123160.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137232.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139869.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156556.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 16-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 23; DB 21; Length 79;
Best Local Similarity 100.0%; Pred. No. 2,6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 KVGIV 5
Db 7 KVGIV 11

RESULT 8
ABG09451
ID ABG09451 standard; Protein; 79 AA.

XX ABG09451;
 XX 13-FEB-2002 (first entry)
 DE Novel human diagnostic protein #9442.
 DE Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS
 XX MO200175067-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US08631.
 PF
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI
 XX WPI: 2001-639362/73.
 DR N-PSDB; AAS73638.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PT
 XX Claim 20; SEQ ID NO 39810; 103pp; English.
 PS
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations in
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 79 AA;
 Query Match 100.0%; Score 23; DB 22; Length 79;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 16483.
 DE Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS
 XX EPI033405-A2.
 PN
 XX 06-SEP-2000.
 PD
 XX 25-FEB-2000; 2000EP-0301439.
 PF
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134321.
 PR 18-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136382.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.

RESULT 9
 AAG16008
 ID AAG16008 standard; Protein; 85 AA.
 XX
 AC AAG16008;

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 23; DB 21; Length 85;
Best Local Similarity 100.0%; Pred No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVCIV 5
|||||
Db 7 KVCIV 11

RESULT 10
ABG18024

ID: ABG18024 standard; Protein; 90 AA.

XX ABG18024;

AC 18-FEB-2002 (first entry)

DT Novel human diagnostic protein #18015.

DE Human; chromosome mapping; gene mapping; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151330.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 23; DB 21; Length 92;
Best Local Similarity 100.0%; Pred. No. 3; 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
DB 7 KVGIV 11

RESULT 12

AAG34700
ID AAG34700 standard; Protein; 92 AA.

XX
AC AAG34700;

XX
DT 18-OCT-2000 (first entry)

XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 42266.

XX
KM Protein identification: signal transduction pathway; metabolic pathway;

KW Hybridisation assay; genetic mapping; gene expression control; promoter;

XX
OS Arabidopsis thaliana.

XX
PN EP1033405-A2.

XX
PD 06-SEP-2000.

XX
PF 25-FEB-2000; 2000EP-0301439.

XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0122180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144337.
PR 19-JUL-1999; 99US-0144338.
PR 19-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145228.
PR 23-JUL-1999; 99US-0145228.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 23; DB 21; Length 92;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
DB 7 KVGIV 11

RESULT 13
ABB61844
ID ABB61844. standard; Protein; 92 AA.

AC ABB61844.

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 12324.

KW Drosophila: developmental biology; cell signalling; insecticide;
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;
XX
XX
XX

DR WPI; 2001-656860/75.
DR N-PSDB; ABL05947.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 12324; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC sequences genomic DNA sequences (AB16176-AB16175), expressed DNA
CC sequences (AB16173-AB16172).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 92 AA;

Query Match 100.0%; Score 23; DB 22; Length 92;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
DB 7 KVGIV 11

RESULT 14
AAB57210
ID AAB57210 standard; Protein; 95 AA.

AC AAB57210;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1788.

KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytosolic; cardiovascular; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.

XX Homo sapiens.

OS Homo sapiens.

PN WO200055174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05988.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.
(ROSE/) ROSEN C A.

PI Rosen CA, Ruben SW;

DR WPI; 2000-587513/55.

DR N-PSDB; AAF16413.

PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
PS Claim 11; Page 2253-2254; 2338pp; English.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 16:00:24 ; Search time 14.5263 Seconds
(without alignments)
12.153 Million cell updates/sec

Title: US-10-054-968-7

Perfect score: 36

Sequence: 1 FGPKNF 6

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCFUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	288	4	US-09-134-001C-3292
2	30	83.3	24	1	US-08-616-368A-40
3	30	83.3	24	1	US-09-054-298A-40
4	30	83.3	24	4	US-08-818-655-40
5	30	83.3	193	1	US-08-616-368A-1
6	30	83.3	193	1	US-08-616-368A-8
7	30	83.3	193	1	US-08-616-368A-9
8	30	83.3	193	1	US-08-616-368A-10
9	30	83.3	193	1	US-08-616-368A-13
10	30	83.3	193	2	US-08-739-485-1
11	30	83.3	193	2	US-08-739-485-9
12	30	83.3	193	2	US-08-739-485-9
13	30	83.3	193	4	US-09-054-298-1
14	30	83.3	193	4	US-09-054-298-8
15	30	83.3	193	4	US-09-054-298-9
16	30	83.3	193	4	US-09-054-298-10
17	30	83.3	193	4	US-09-054-298-13
18	30	83.3	193	4	US-08-818-655-1
19	30	83.3	193	4	US-08-818-655-8
20	30	83.3	193	4	US-08-818-655-9
21	30	83.3	193	4	US-08-818-655-10
22	30	83.3	193	4	US-08-818-655-13
23	30	83.3	193	4	US-08-818-655-10
24	30	83.3	806	4	US-08-945-983-2
25	30	83.3	941	4	US-09-513-783A-172
26	30	83.3	1024	4	US-09-562-737-50
27	30	83.3	2756	1	US-08-375-709-11

28	30	83.3	2756	1	US-08-752-929-11	Sequence 11, Appl
29	30	83.3	2756	4	US-09-090-793-7	Sequence 7, Appl
30	28	77.8	576	3	US-08-948-564-16	Sequence 16, Appl
31	28	77.8	1074	1	US-07-912-015-4	Sequence 4, Appl
32	27	75.0	50	5	PCT-US91-02942-56	Sequence 56, Appl
33	27	75.0	117	1	US-07-634-278-4	Sequence 4, Appl
34	27	75.0	117	1	US-07-634-278-15	Sequence 15, Appl
35	27	75.0	117	1	US-07-634-278-12	Sequence 12, Appl
36	27	75.0	117	1	US-07-634-278-104	Sequence 104, Appl
37	27	75.0	117	1	US-08-477-728-4	Sequence 4, Appl
38	27	75.0	117	1	US-08-477-728-15	Sequence 15, Appl
39	27	75.0	117	1	US-08-477-728-72	Sequence 72, Appl
40	27	75.0	117	1	US-08-477-728-104	Sequence 104, Appl
41	27	75.0	117	1	US-08-474-040-4	Sequence 4, Appl
42	27	75.0	117	1	US-08-474-040-15	Sequence 15, Appl
43	27	75.0	117	1	US-08-474-040-72	Sequence 72, Appl
44	27	75.0	117	1	US-08-474-040-104	Sequence 104, Appl
45	27	75.0	117	1	US-08-487-200-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-134-001C-3292
Sequence 3292, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GNC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13, 964
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3292
LENGTH: 288
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3292

Query Match 100.0%; Score 36; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FGPKNF 6
Db 111 FGPKNF 116

RESULT 2
US-08-616-368A-40
Sequence 40, Application US/08616368A
Patent No. 5767262
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Haber, Edgar
APPLICANT: Jain, Mukesh
APPLICANT: Yet, Shaw-Fang
TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,368A
FILING DATE: 15-MAR-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-616-368A-40

Query Match 83.3%; Score 30; DB 1; Length 24;
Best Local Similarity 83.3%; Pred. No. 7.1;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
|||||
DB 6 FGPKGF 11

RESULT 3
US-09-054-298-40
Sequence 40, Application US/09054298
Patent No. 6136953
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Haber, Edgar
APPLICANT: Jain, Mukesh
APPLICANT: Yet, Shaw-Fang
TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,298
FILING DATE: 02-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/616,368
FILING DATE: 15-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Beattie, Ph.D., Ingrid A.
REGISTRATION NUMBER: P-42,306
REFERENCE/DOCKET NUMBER: 05433/022002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:

LENGTH: 24 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-054-298-40

Query Match 83.3%; Score 30; DB 4; Length 24;
Best Local Similarity 83.3%; Pred. No. 7.1;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
|||||
DB 6 FGPKGF 11

RESULT 4
US-08-818-655-40
Sequence 40, Application US/08818655
Patent No. 6258557
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Haber, Edgar
APPLICANT: Jain, Mukesh
APPLICANT: Yet, Shaw-Fang
TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,655
FILING DATE: 14-MAR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/616,368
FILING DATE: 15-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/030001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-818-655-40

Query Match 83.3%; Score 30; DB 4; Length 24;
Best Local Similarity 83.3%; Pred. No. 7.1;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
|||||
DB 6 FGPKGF 11

RESULT 5
US-08-616-368A-1
Sequence 1, Application US/08616368A
Patent No. 5767262

GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Haber, Edgar
APPLICANT: Jain, Mukesh
APPLICANT: Yet, Shaw-Fang
TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,368A
FILING DATE: 15-MAR-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-616-368A-1

Query Match 83.3%; Score 30; DB 1; Length 193;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FGPKNF 6
Db 175 FGPKNF 180

RESULT 6
US-08-616-368A-8
Sequence 8, Application US/08616368A
Patent No. 5767262
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Haber, Edgar
APPLICANT: Jain, Mukesh
APPLICANT: Yet, Shaw-Fang
TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/616,368A
FILING DATE: 15-MAR-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-616-368A-8

Query Match 83.3%; Score 30; DB 1; Length 193;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FGPKNF 6
Db 175 FGPKNF 180

RESULT 7
US-08-616-368A-9
Sequence 9, Application US/08616368A
Patent No. 5767262
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Haber, Edgar
APPLICANT: Jain, Mukesh
APPLICANT: Yet, Shaw-Fang
TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,368A
FILING DATE: 15-MAR-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-616-368A-9

Query Match 83.3%; Score 30; DB 1; Length 193;

Best: Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FGPKNF 6
|||||
Db 175 FGPKGF 180

RESULT 8

US-08-616-368A-10
; Sequence 10, Application US/08616368A
; Patent No. 5767262
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Haber, Edgar
; APPLICANT: Jain, Mukesh
; APPLICANT: Yet, Shaw-Fang
; TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616,368A
; FILING DATE: 15-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 193 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-616-368A-10

Query Match 83.3%; Score 30; DB 1; Length 193;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FGPKNF 6
|||||
Db 175 FGPKGF 180

RESULT 9

US-08-616-368A-13
; Sequence 13, Application US/08616368A
; Patent No. 5767262
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Haber, Edgar
; APPLICANT: Jain, Mukesh
; APPLICANT: Yet, Shaw-Fang
; TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,368A
FILING DATE: 15-MAR-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-616-368A-13

Query Match 83.3%; Score 30; DB 1; Length 193;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FGPKNF 6
|||||
Db 175 FGPKGF 180

RESULT 10

US-08-739-485-1
; Sequence 1, Application US/08739485
; Patent No. 5863898
; GENERAL INFORMATION:
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: NOVEL HUMAN LIM PROTEINS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/739,485
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0142 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
US-08-739-485-1

Query Match 83.3%; Score 30; DB 2; Length 193;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
Db 175 FGPKGF 180

RESULT 11
US-08-739-485-7
Sequence 7, Application US/08739485
Patent No. 5863898
GENERAL INFORMATION:
APPLICANT: Goli, Surya K.
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
TITLE OF INVENTION: NOVEL HUMAN LIM PROTEINS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/739,485
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0142 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 118161
US-08-739-485-7

Query Match 83.3%; Score 30; DB 2; Length 193;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
Db 175 FGPKGF 180

RESULT 12
US-08-739-485-9
Sequence 9, Application US/08739485
Patent No. 5863898
GENERAL INFORMATION:
APPLICANT: Goli, Surya K.
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
TITLE OF INVENTION: NOVEL HUMAN LIM PROTEINS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/739,485
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0142 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1314351
US-08-739-485-9

Query Match 83.3%; Score 30; DB 2; Length 193;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
Db 175 FGPKGF 180

RESULT 13
US-09-054-298-1
Sequence 1, Application US/09054298
Patent No. 6136953
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Haber, Edgar

APPLICANT: Jain, Mukesh
APPLICANT: Yet, Shaw-Fang
TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,298
FILING DATE: 02-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/616,368
FILING DATE: 15-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Beattie, Ph.D., Ingrid A.
REGISTRATION NUMBER: P-42,306
REFERENCE/DOCKET NUMBER: 05433/022002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-054-298-1

Query Match 83.3%; Score 30; DB 4; Length 193;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
DB 175 FGPKNF 180

RESULT 14
US-09-054-298-8
Sequence 8, Application US/09054298
Patent No. 6136953
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Haber, Edgar
APPLICANT: Jain, Mukesh
APPLICANT: Yet, Shaw-Fang
TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,298
FILING DATE: 02-APR-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/616,368
FILING DATE: 15-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Beattie, Ph.D., Ingrid A.
REGISTRATION NUMBER: P-42,306
REFERENCE/DOCKET NUMBER: 05433/022002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-09-054-298-8

Query Match 83.3%; Score 30; DB 4; Length 193;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
DB 175 FGPKNF 180

RESULT 15
US-09-054-298-9
Sequence 9, Application US/09054298
Patent No. 6136953
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Haber, Edgar
APPLICANT: Jain, Mukesh
APPLICANT: Yet, Shaw-Fang
TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,298
FILING DATE: 02-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/616,368
FILING DATE: 15-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Beattie, Ph.D., Ingrid A.
REGISTRATION NUMBER: P-42,306
REFERENCE/DOCKET NUMBER: 05433/022002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-054-298-9

Query Match 83.3%; Score 30; DB 4; Length 193;
 Best Local Similarity 83.3%; Pred. No. 54;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FGPKNF 6
 |||||
 Db 175 FGPKNF 180

Search completed: July 18, 2003, 16:05:18
 Job time : 15.5263 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 16:00:24 ; Search time 12.1053 Seconds
(without alignments)
12.153 Million cell updates/sec

Title: US-10-054-968-6
Perfect score: 23
Sequence: 1 KVGIV 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-AA: *
1: /cgn2-6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2-6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2-6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2-6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2-6/ptodata/1/1aa/PCUTUS.COMB.pep: *
6: /cgn2-6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	100.0	116	4	US-09-071-035-438
2	23	100.0	228	3	US-08-971-158-5
3	23	100.0	228	4	US-09-416-488-5
4	23	100.0	269	2	US-08-978-404B-10
5	23	100.0	496	1	US-08-462-128-37
6	23	100.0	496	1	US-08-463-160-37
7	23	100.0	496	2	US-08-001-078A-1
8	23	100.0	496	2	US-08-897-443-4
9	23	100.0	496	2	US-08-463-218-1
10	23	100.0	496	5	PCT-US94-00253-1
11	23	100.0	707	4	US-09-021-560-4
12	23	100.0	744	2	US-08-462-080B-2
13	23	100.0	744	3	US-08-462-090-2
14	23	100.0	744	3	US-08-463-461-2
15	23	100.0	744	4	US-09-021-560-2
16	23	100.0	1228	4	US-09-605-785-537
17	23	100.0	1228	4	US-09-439-313-537
18	23	100.0	1261	4	US-09-605-785-538
19	23	100.0	1261	4	US-09-439-313-538
20	23	100.0	1528	1	US-08-463-092B-6
21	23	100.0	1528	2	US-08-462-109A-6
22	23	100.0	1528	2	US-08-460-907B-6
23	23	100.0	1528	3	US-08-463-179A-6
24	23	100.0	1528	3	US-08-461-384B-6
25	23	100.0	1531	1	US-08-141-893-2
26	23	100.0	1531	1	US-08-463-092B-2
27	23	100.0	1531	1	US-08-463-092B-4

28	23	100.0	1531	2	US-08-462-109A-2	Sequence 2, Appli
29	23	100.0	1531	2	US-08-462-109A-4	Sequence 4, Appli
30	23	100.0	1531	2	US-08-460-907B-2	Sequence 2, Appli
31	23	100.0	1531	2	US-08-460-907B-4	Sequence 4, Appli
32	23	100.0	1531	3	US-08-463-179A-2	Sequence 2, Appli
33	23	100.0	1531	3	US-08-463-179A-4	Sequence 4, Appli
34	23	100.0	1531	3	US-08-461-384B-2	Sequence 2, Appli
35	23	100.0	1531	3	US-08-461-384B-4	Sequence 4, Appli
36	23	100.0	1531	3	US-08-407-207A-2	Sequence 2, Appli
37	23	100.0	1548	1	US-08-463-092B-7	Sequence 7, Appli
38	23	100.0	1548	2	US-08-460-907B-7	Sequence 7, Appli
39	23	100.0	1621	4	US-08-972-927-3	Sequence 3, Appli
40	23	100.0	1622	4	US-08-972-927-6	Sequence 6, Appli
41	22	95.7	29	1	US-07-883-949B-5	Sequence 5, Appli
42	22	95.7	29	1	US-07-883-949B-6	Sequence 6, Appli
43	22	95.7	50	4	US-09-447-125B-16	Sequence 16, Appli
44	22	95.7	50	4	US-09-447-125B-17	Sequence 17, Appli
45	22	95.7	50	4	US-09-447-125B-18	Sequence 18, Appli

ALIGNMENTS

RESULT 1
US-09-071-035-438
; Sequence 438, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brooks
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 438:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 116 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-071-035-438
Query Match 100.0%; Score 23; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KVGIV 5
DB 5 KVGIV 9

RESULT 2
US-08-971-158-5
; Sequence 5, Application US/08971158
; Patent No. 6010879
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: HUMAN MITOCHONDRIAL CHAPERONE PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971.158
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/824,875
; FILING DATE: March 26, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0245-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 493576
; US-08-971-158-5
Query Match 100.0%; Score 23; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVGIV 5
Db 219 KVGIV 223
RESULT 3
US-09-416-488-5
; Sequence 5, Application US/09416488
; Patent No. 6432915
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: HUMAN MITOCHONDRIAL CHAPERONE PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA

ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/416.488
; FILING DATE: 12-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/971.158
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0245-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 493576
; US-09-416-488-5
Query Match 100.0%; Score 23; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVGIV 5
Db 219 KVGIV 223
RESULT 4
US-08-978-404B-10
; Sequence 10, Application US/08978404B
; Patent No. 5968782
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAST CELL PROTEINASE THAT CLEAVES
; TITLE OF INVENTION: FIBRINOGEN
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978.404B
; FILING DATE: 25-NOV-97
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/032,354
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637

REFERENCE/DOCKET NUMBER: B0801/7090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 269 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5968782e
US-08-978-404B-10

Query Match 100.0%; Score 23; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
|||||
DB 17 KVGIV 21

RESULT 5
US-08-462-128-37
Sequence 37, Application US/08462128
Patent No. 5686059
GENERAL INFORMATION:
APPLICANT: Goelink, Paul F.
APPLICANT: Tondravil, Mehdad
TITLE OF INVENTION: CARTILAGE MATRIX PROTEIN AND METHODS FOR
TITLE OF INVENTION: USE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,128
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/006,096
FILING DATE: 15-JAN-1993
APPLICATION NUMBER: US 07/866,403
FILING DATE: 10-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-462-128-37

Query Match 100.0%; Score 23; DB 1; Length 496;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KVGIV 5

DB |||||
378 KVGIV 382

RESULT 6
US-08-463-180-37
Sequence 37, Application US/08463180
Patent No. 5741670
GENERAL INFORMATION:
APPLICANT: Goelink, Paul F.
APPLICANT: Tondravil, Mehdad
TITLE OF INVENTION: CARTILAGE MATRIX PROTEIN AND METHODS FOR
TITLE OF INVENTION: USE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,180
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/006,096
FILING DATE: 15-JAN-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/866,403
FILING DATE: 10-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-005DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-463-180-37

Query Match 100.0%; Score 23; DB 1; Length 496;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
|||||
DB 378 KVGIV 382

RESULT 7
US-08-001-078A-1
Sequence 1, Application US/08001078A
Patent No. 5872094
GENERAL INFORMATION:
APPLICANT: Goelink, Paul F.
APPLICANT: Tondravil, M., Mehdad
APPLICANT: Binette, Francois
TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
TITLE OF INVENTION: FORMATION
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 STATE STREET, Suite 510
CITY: BOSTON
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/001,078A
FILING DATE: 06-JAN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-001-078A-1

Query Match 100.0%; Score 23; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
DB 378 KVGIV 382

RESULT 8
US-08-897-443-4
Sequence 4, Application US/08897443
Patent No. 5981263
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
APPLICANT: Kaser, Mathew
TITLE OF INVENTION: HUMAN MATRILIN-3
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,443
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0348 US
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1732121
US-08-897-443-4

Query Match 100.0%; Score 23; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
DB 378 KVGIV 382

RESULT 9
US-08-463-218-1
Sequence 1, Application US/08463218
Patent No. 5986052
GENERAL INFORMATION:
APPLICANT: Goelink, Paul F.
APPLICANT: Tondravil, Mehrdad
APPLICANT: Binette, Francois
TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
FORMATION
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, Suite 510
CITY: BOSTON
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,218
FILING DATE: 05-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,078
FILING DATE: 06-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-008DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-463-218-1

Query Match 100.0%; Score 23; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
DB 378 KVGIV 382

RESULT 10
PCT-US94-00253-1
Sequence 1, Application PC/TUS9400253
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
TITLE OF INVENTION: FORMATION
NUMBER OF SEQUENCES: 2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00253
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,078
FILING DATE: 06-JAN-1993
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US94-00253-1

Query Match 100.0%; Score 23; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGVIV 5
DB 378 KGVIV 382

RESULT 11
US-09-021-560-4
Sequence 4, Application US/09021560
Patent No. 6410719
GENERAL INFORMATION:
APPLICANT: BOREN, THOMAS
APPLICANT: NORMARK, STAFFAN
APPLICANT: ARNOVIST, ANNA
APPLICANT: LIVER, DAG
TITLE OF INVENTION: BLOOD GROUP ANTIGEN BINDING PROTEIN AND
TITLE OF INVENTION: CORRESPONDING GENES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,560
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 825-144P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 707 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-021-560-4

Query Match 100.0%; Score 23; DB 4; Length 707;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGVIV 5
DB 517 KGVIV 521

RESULT 12
US-08-462-080B-2
Sequence 2, Application US/08462080B
Patent No. 5997913
GENERAL INFORMATION:
APPLICANT: Fowler, Timothy
APPLICANT: Barnett, Christopher C.
APPLICANT: Shoemaker, Sharon
TITLE OF INVENTION: Saccharification of Cellulose by Cloning and
TITLE OF INVENTION: Amplification of the Beta-glucosidase Gene of Trichoderma R
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: Ca
COUNTRY: U.S.A.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462;080B
FILING DATE: 05-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/248,586
FILING DATE: 24-MAY-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/807,028
FILING DATE: 10-DEC-1991
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/625,140
FILING DATE: 10-DEC-1990
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Stone, Christopher L.
REGISTRATION NUMBER: 35,696
REFERENCE/DOCKET NUMBER: GC78D3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-846-7555
TELEFAX: 650-845-6504
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 744 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-080B-2

Query Match 100.0%; Score 23; DB 2; Length 744;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGVIV 5
|||||
Db 59 KGVIV 63

RESULT 13
US-08-462-090-2
; Sequence 2, Application US/08462090
; Patent No. 6022725
GENERAL INFORMATION:
APPLICANT: Fowler, Timothy
APPLICANT: Barnett, Christopher C.
APPLICANT: Shoemaker, Sharon
TITLE OF INVENTION: Saccharification of Cellulose by Cloning
TITLE OF INVENTION: and Amplification of the Beta-glucosidase Gene of
TITLE OF INVENTION: Trichoderma Reesel
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Building, 699 Prince St.
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,090
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/625,140
FILING DATE: 10-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Dillahunty, T. Gene
REGISTRATION NUMBER: 25,423
REFERENCE/DOCKET NUMBER: 010055-056
TELEPHONE: 415-854-7400
TELEFAX: 415-854-8275
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 744 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-090-2

Query Match 100.0%; Score 23; DB 3; Length 744;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGVIV 5
|||||
Db 59 KGVIV 63

RESULT 14
US-08-463-461-2
; Sequence 2, Application US/08463461
; Patent No. 6103464
GENERAL INFORMATION:
APPLICANT: Fowler, Timothy
APPLICANT: Barnett, Christopher C.
APPLICANT: Shoemaker, Sharon
TITLE OF INVENTION: Saccharification of Cellulose by Cloning
TITLE OF INVENTION: and Amplification of the Beta-glucosidase Gene of
TITLE OF INVENTION: Trichoderma Reesel

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: U.S.A.
ZIP: 94304-1013
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,461
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Christopher L. Stone
REGISTRATION NUMBER: 35,696
REFERENCE/DOCKET NUMBER: GC78D4
TELEPHONE: 415-846-7555
TELEFAX: 415-845-6504
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 744 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-461-2

Query Match 100.0%; Score 23; DB 3; Length 744;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGVIV 5
|||||
Db 59 KGVIV 63

RESULT 15
US-09-021-560-2
; Sequence 2, Application US/09021560
; Patent No. 6410719
GENERAL INFORMATION:
APPLICANT: BOREN, THOMAS
APPLICANT: NORMARK, STAFFAN
APPLICANT: ARNOVIST, ANNA
APPLICANT: ILVER, DAG
TITLE OF INVENTION: BLOOD GROUP ANTIGEN BINDING PROTEIN AND
TITLE OF INVENTION: CORRESPONDING GENES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,560
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 825-144P

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 744 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-021-560-2

Query Match 100.0%; Score 23; DB 4; Length 744;
 Best local Similarity 100.0%; Pred. No. 6.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
 |||||
 Db 554 KVGIV 558

Search completed: July 18, 2003, 16:05:17
 Job time : 13.1053 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:57:45 : Search time 5.52632 Seconds
(Without alignments)
37.526 Million cell updates/sec

Title: US-10-054-968-6
Perfect score: 23
Sequence: 1 KVGIV 5

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	100.0	52	INS_ACIQU	P01423 acipenser g
2	23	100.0	52	INS_LEPSP	P09476 lepisosteus
3	23	100.0	84	R17_ODOSI	P49504 odontella s
4	23	100.0	90	R37A_OSTOS	O61598 osteria g
5	23	100.0	91	R37A_CHICK	P32046 gallus gall
6	23	100.0	91	R37A_CHICK	O61462 cyprochito
7	23	100.0	91	R37A_HUMAN	P12751 homo sapien
8	23	100.0	91	R37A_HUMAN	O90470 ictalurus p
9	23	100.0	91	R37A_ICTPU	O17307 schistosoma
10	23	100.0	92	R37A_SCHMA	P43209 brassica ra
11	23	100.0	156	R1SB_BACHD	O8584 methanococ
12	23	100.0	156	R1SB_BACHD	O8584 methanococ
13	23	100.0	177	R10_THETN	O87704 thermoaer
14	23	100.0	220	Y132_METJA	O87596 methanococ
15	23	100.0	228	GRPE_YEAST	P38523 saccharomyc
16	23	100.0	254	YIGE_ECOLI	P27840 escherichia
17	23	100.0	269	TRYM_CAMEA	P19236 canis fami
18	23	100.0	280	FRHB_METVO	O00391 methanococ
19	23	100.0	288	SUCD_ECOLI	P07459 escherichia
20	23	100.0	291	Y263_MYCGE	P45102 haemophilus
21	23	100.0	293	SUCD_HAEN	P33591 coxiella bu
22	23	100.0	294	SUCD_COXAB	O15567 pseudomonas
23	23	100.0	295	SUCD_PSRAB	P13735 thermus agu
24	23	100.0	310	LDH_THEAO	P06150 thermus cal
25	23	100.0	310	LDH_THEAO	O50734 mycobacteri
26	23	100.0	330	YP64_MYCTU	O9K872 bacillus ha
27	23	100.0	345	ARGC_BACHD	P00805 escherichia
28	23	100.0	348	ASG2_ECOLI	P26546 human papil
29	23	100.0	358	VE2_HPV58	O4894 saccharomyc
30	23	100.0	360	YM97_YEAST	O75936 homo sapien
31	23	100.0	387	BODG_HUMAN	O924Y0 mus musculu
32	23	100.0	387	BODG_MOUSE	O924Y0 mus musculu
33	23	100.0	387	BODG_MOUSE	O924Y0 mus musculu

34	23	100.0	388	1	AMPC_YEREN	P45460 yersinia en
35	23	100.0	409	1	MAS5_YEAST	P25491 saccharomyc
36	23	100.0	414	1	RCA_AMASP	P58555 anabaena sp
37	23	100.0	415	1	RCA_AMASP	O06721 anabaena sp
38	23	100.0	464	1	GATH_YEAST	O03557 saccharomyc
39	23	100.0	493	1	CAMA_CHICK	P05099 gallus galli
40	23	100.0	496	1	CAMA_HUMAN	P21941 homo sapien
41	23	100.0	500	1	CAMA_MOUSE	P51942 mus musculu
42	23	100.0	504	1	AMID_SULSO	P95896 sulfobius
43	23	100.0	521	1	TSAG_RICIS	P37917 rickettsia
44	23	100.0	524	1	TSAG_RICIS	P22940 rickettsia
45	23	100.0	656	1	FEMI_CAEBL	P17221 caenorhabdi

ALIGNMENTS

RESULT 1	INS_ACIQU	STANDARD:	PRT:	52 AA.
AC	P01423:			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Insulin.			
GN	INS.			
OS	Acipenser guldenstadti (Caspian sturgeon) (Russian sturgeon).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;			
OC	Acipenser.			
OX	NCBI_TaxID=7902;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Pancreas;			
RC	MEDLINE=98312733; PubMed=9650713;			
RA	Rusakov V.I., Moriyama S., Bondareva V.M., Kolychiev A.P., Amemiya Y.,			
RA	Yasuda A., Kawachi H.;			
RT	"Isolation and characterization of insulin in Russian sturgeon			
RT	(Acipenser guldenstadti).";			
RL	J. Pept. Res. 51:395-400(1998).			
CC	-1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT			
CC	INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND			
CC	FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE			
CC	CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.			
CC	-1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO			
CC	DISULFIDE BONDS.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.			
DR	HSSP; P01308; ISEN.			
DR	InterPro: IPR004825; Ins/IGF/relax.			
DR	PRINTS: PR00276; INSULINB.			
DR	PRINTS: PR00277; INSULINA.			
DR	SMART; SM00078; IIGF. 1.			
DR	PROSITE; PS00262; INSULIN; 1.			
FT	Insulin family; Hormone; Glucose metabolism.			
FT	CHAIN	1	31	INSULIN B CHAIN.
FT	NON_CONS	31	32	
FT	CHAIN	32	52	INSULIN A CHAIN.
FT	DISULFID	7	38	INTERCHAIN.
FT	DISULFID	19	51	INTERCHAIN.
FT	DISULFID	37	42	
SO	SEQUENCE	52 AA;	5811 MW;	BDID693998041631 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 52;
Best Local Similarity 100.0%; Pred. No. 32;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
DB 30 KVGIV 34

RESULT 2

```

INS_LEPSP STANDARD; PRT; 52 AA.
ID INS_LEPSP
AC P09476;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Insulin.
GN INS.
OS Lepisosteus spatula (Alligator gar) (Atractosteus spatula).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Semionotiformes; Lepisosteidae;
OC Lepisosteus
OX NCBI_TaxID=917;
RN [1]
RP SEQUENCE.
RX MEDLINE=88030594; PubMed=3311873;
RA Pollock H.G., Kimmel J.R., Hamilton J.W., Rouse J.B., Ebner K.E.,
RA Lance V., Rawlitch A.B.;
RT Isolation and structures of alligator gar (Lepisosteus spatula)
RT insulin and pancreatic polypeptide."
RL Gen. Comp. Endocrinol. 67:375-382(1987).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR HSSP; P01308; IIGX.
DR InterPro; IPR004825; Ins/IGF/relax.
DR PRINTS; PR00276; INSULINA.
DR SMART; SM00078; ILGF. 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism.
FT CHAIN 1 31 INSULIN B CHAIN.
FT NON_CONS 31 32
FT CHAIN 32 52 INSULIN A CHAIN.
FT DISULFID 7 38 INTERCHAIN.
FT DISULFID 19 51 INTERCHAIN.
FT DISULFID 37 42
SQ SEQUENCE 52 AA; 5865 MW; 6D61138386AC0F10 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 52;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
DB 30 KVGIV 34

RESULT 3
RRI7_ODOSI STANDARD; PRT; 84 AA.
ID RRI7_ODOSI
AC P49504;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Chloroplast 30S ribosomal protein S17.
GN RPS17.
OS Odontella sinensis (Marine centric diatom).
OC Chloroplast.
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Bidulphiophycidae; Eupodiscaceae; Odontella.
OX NCBI_TaxID=2839;
RN [1]
RP SEQUENCE FROM N.A.
RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
Odontella sinensis.";
```

```

RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -1- SIMILARITY: BELONGS TO THE S17P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; Z67753; CAA91639.1; -.
DR HSSP; P23828; IIRP.
DR InterPro; IPR000266; Ribosomal_S17.
DR Pfam; PF00366; Ribosomal_S17; 1.
DR PRINTS; PR00973; RIBOSOMAL_S17.
DR ProDom; PD001295; Ribosomal_S17; 1.
DR PROSITE; PS00056; RIBOSOMAL_S17; 1.
KW Ribosomal protein; Chloroplast.
SQ SEQUENCE 84 AA; 9638 MW; 4B86C2C61C2FBD02 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
DB 6 KVGIV 10

RESULT 4
R37A_OSTOS STANDARD; PRT; 90 AA.
ID R37A_OSTOS
AC O61598;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S ribosomal protein L37a.
GN RPL37A.
OS Ostertagia ostertagi.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemonchidae; Ostertagiinae; Ostertagia.
OX NCBI_TaxID=6317;
RN [1]
RP SEQUENCE FROM N.A.
RA Moore J., Devaney E.;
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L37AE FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AF052737; AAC08431.1; -.
DR InterPro; IPR002674; Ribosomal_L37ae.
DR Pfam; PF01780; Ribosomal_L37ae; 1.
DR ProDom; PD006591; Ribosomal_L37ae; 1.
DR TIGRPFAMS; TIGR00280; L37a; 1.
KW Ribosomal protein; Zinc-finger; Metal-binding.
FT INIT_MET 0 0 BY SIMILARITY.
FT ZN_FING 38 59 CA-TYPE.
SQ SEQUENCE 90 AA; 9977 MW; FCA925488B5D0471 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
DB 1 KVGIV 5
```



```

RT differential gene expression in tissues and cell lines."
RL gene 132:285-289(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Brain, Skin, and Testis;
RA Strausberg R.;
RL submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=Rat; STRAIN=Sprague-Dawley;
RA MEDLINE=89325328; PubMed=2546769;
RA Tanaka T., Aoyama Y., Chan Y.-L., Wool I.G.;
RA "The primary structure of rat ribosomal protein L37a.";
RL Eur. J. Biochem. 183:15-18(1989).
RN [6]
RP SEQUENCE OF 1-28.
RC SPECIES=Rat;
RA MEDLINE=80252792; PubMed=398910;
RA Wittmann-Jiebold B., Geissler A.W., Lin A., Wool I.G.;
RA "Sequence of the amino-terminal region of rat liver ribosomal
RA proteins S4, S6, S8, L6, L7a, L18, L27, L30, L37, L37a, and L39.";
RL J. Supramol. Struct. 12:425-433(1979).
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse; STRAIN=C57BL/6; TISSUE=Spleen;
RA MEDLINE=94021399; PubMed=8415004;
RA Su Y., Babu N., Raj K., Au W.-C., Pitcha P.M.;
RA "Primary sequence of the mouse ribosomal protein L37a.";
RL Nucleic Acids Res. 21:4400-4400(1993).
CC -1- SIMILARITY: BELONGS TO THE L37AE FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL; L06499; AAA60280.1; -
CC DR EMBL; X66699; CAA47244.1; -
CC DR EMBL; L22154; -: NOT ANNOTATED_CDS.
CC DR EMBL; BC000555; AAH00555.1; -
CC DR EMBL; BC014262; AAH16748.1; -
CC DR EMBL; BC016748; AAH16748.1; -
CC DR EMBL; X14069; CAA32232.1; -
CC DR EMBL; X73331; CAA51758.1; -
CC DR PIR; S05014; R5RT37.
CC DR PIR; JN0875; JN0875.
CC DR PIR; S26379; S26379.
CC DR PIR; S34284; S34284.
CC DR PIR; S42109; S42109.
CC DR Genew; HGNC:10348; RPL37A.
CC DR MGD; MGI:98068; Rpl37a.
CC DR InterPro; IPR002674; Ribosomal_L37ae.
CC DR Pfam; PF01780; Ribosomal_L37ae; 1.
CC DR ProDom; PD006591; Ribosomal_L37ae; 1.
CC DR TIGRfams; TIGR00280; L37a; 1.
CC DR Ribosomal protein; Zinc-finger; Metal-binding.
CC FT INIT_MET 0
CC FT ZN_FING 38
CC FT 59
CC SEQUENCE 91 AA; 10144 MW; 11DB5FCE2DAB986B CRC64;

```

```

R37A_ICTPU STANDARD: PRT: 91 AA.
ID R37A_ICTPU
AC Q90YT0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L37a.
GN RPL37A.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
OX NCBI_TaxID=7998;
RN [1]
RP SEQUENCE FROM N.A.
RA Patterson A.P., Karsi A., Liu Z.J.;
RT "Translational machinery of channel catfish: II. Complementary DNA and
RT expression of the complete set of 47 60S ribosomal proteins.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L37AE FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF401594; AAK95166.1;
DR InterPro: IPR002674; Ribosomal_L37ae.
DR Pfam: PF01780; Ribosomal_L37ae; 1.
DR ProDom: PD006591; Ribosomal_L37ae; 1.
DR TRIGRAMS: TIGR00280; L37a; 1.
KW Ribosomal protein; zinc-finger; Metal-binding.
FT INIT MET 0
FT ZN FING 38 59
FT BY SIMILARITY.
FT C4-TYPE.
SO SEQUENCE 91 AA; 10134 MW; 0EE65FCF9DAB986B CRC64;

Query Match 100.0%; Score 23; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY: 1 KVGIV 5
D6: 6 KVGIV 10

RESULT 9
R37A_SCHMA STANDARD: PRT: 91 AA.
ID R37A_SCHMA
AC O17307;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S ribosomal protein L37a (Fragment).
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeiida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-NMRI;
RA Razaq M.M., Shalaby K.A., Mohamed M.M., Karim A.M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L37AE FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: AF025539; AAB81969.1;
DR InterPro: IPR002674; Ribosomal_L37ae.
DR Pfam: PF01780; Ribosomal_L37ae.1.
DR ProDom: PD006591; Ribosomal_L37ae; 1.
DR TIGRFAMs: TIGR00280; L37a; 1.
KW Ribosomal protein; zinc-finger; Metal-binding.
FT NON_TER 1
FT ZN_FING 38
SQ SEQUENCE 91 AA; 10005 MW; 46109EBC416EA9B0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 1; Length 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
Db 6 KVGIV 10

RESULT 10
R37A_BRARA STANDARD; PRT; 92 AA.
AC P43209;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 60S-ribosomal protein L37a.
GN RPL37A.
OS Brassica rapa (Turnip).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=51350;
RN [1]
RP SEQUENCE FROM N.A.
RA Song S., Choi Y.;
RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L37AE FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.lsb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: Z24739; CAA80864.1;
DR EMBL: L21897; AAA51421.1;
DR InterPro: IPR002674; Ribosomal_L37ae.
DR Pfam: PF01780; Ribosomal_L37ae; 1.
DR ProDom: PD006591; Ribosomal_L37ae; 1.
DR TIGRFAMs: TIGR00280; L37a; 1.
KW Ribosomal protein; zinc-finger; Metal-binding.
FT INIT_MET 0
FT ZN_FING 38
SQ SEQUENCE 92 AA; 10247 MW; 9F6D45B9AE2157A6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 1; Length 92;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
Db 6 KVGIV 10

RESULT 11
RISB_BACHD STANDARD; PRT; 156 AA.
AC G9KCL4;

```

```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9) (DMRL synthase)
DE (lumazine synthase) (Riboflavin synthase beta chain).
GN RIBH OR BH157.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=96655;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex
CC catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-
CC ribityl-amino-2,4(1H,3H)-pyrimidinedione and L-3,4-dihydroxy-2-
CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit
CC catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-amino-
CC 2,4(1H,3H)-pyrimidinedione with L-3,4-dihydroxy-2-butanone-4-
CC phosphate yielding 6,7-dimethyl-8-lumazine (By similarity).
CC -1- CATALYTIC ACTIVITY: 2,6,7-dimethyl-8-(1-D-ribityl)lumazine +
CC riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.
CC -1- PATHWAY: Riboflavin biosynthesis; last step.
CC -1- SIMILARITY: BELONGS TO THE DMRL SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.lsb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: AP001512; BAB05276.1;
DR HSSP: P11998; 1RVV.
DR InterPro: IPR002180; DMRL_synthase.
DR Pfam: PF00885; DMRL_synthase; 1.
DR ProDom: PD003664; DMRL_synthase; 1.
DR TIGRFAMs: TIGR00114; ribH; 1.
KW Riboflavin biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 156 AA; 16353 MW; 5A889472A0B09AB8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 1; Length 156;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
Db 15 KVGIV 19

RESULT 12
RISC_METJA STANDARD; PRT; 156 AA.
ID RISC_METJA
AC Q58584;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Riboflavin synthase (EC 2.5.1.9).
DE RIBH OR Mj1184.
GN RIBH OR Mj1184.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;

```

```

RX MEDLINE-96337999; PubMed-8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1-D-ribityl)lumazine =
CC riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxyprimidine.
CC -1- PATHWAY: Riboflavin biosynthesis; last step.
CC -1- SIMILARITY: BELONGS TO THE DMRL SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: U67559; AAB99185.1; -.
DR TIGR: MJ1184; -.
DR InterPro: IPR002180; DMRL_synthase.
DR Pfam: PF00885; DMRL_synthase; 1.
KM Riboflavin biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 156 AA; 17494 MW; 6D65112C6B1320F5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 1; Length 156;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGVIV 5
Db 4 KGVIV 8

RESULT 13
RL10_THETN
ID RL10_THETN STANDARD; PRT; 177 AA.
AC Q8R704;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L10.
GN RPLJ OR TTE2304.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID-119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MB4 / JCM 11007;
RX MEDLINE-21992816; PubMed-11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----

```

```

DR EMBL: AE013173; AAM25445.1; -.
DR PROSITE: PS01109; RIBOSOMAL_L10; FALSE_NEG.
KM Ribosomal protein; Complete proteome.
SQ SEQUENCE 177 AA; 19591 MW; 50DDCF96EF6F4E8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 1; Length 177;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGVIV 5
Db 110 KGVIV 114

RESULT 14
Y132_METUA
ID Y132_METUA STANDARD; PRT; 220 AA.
AC O57586;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0132.
GN MJ0132.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID-2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE-96337999; PubMed-8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: TO M.JANNASCHII MJ1220 AND MJE142.
CC -1- SIMILARITY: WITH TYPE I RESTRICTION SYSTEM ADENINE METHYLASES
CC (M SUBUNIT).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: U67470; AAB98113.1; -.
DR TIGR: MJ0132; -.
DR InterPro: IPR003356; N6_DNA_Mtase.
DR Pfam: PF02384; N6_Mtase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 220 AA; 25766 MW; 710DDAEC4CA7954 CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 1; Length 220;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGVIV 5
Db 48 KGVIV 52

RESULT 15
GRPE_YEAST
ID GRPE_YEAST STANDARD; PRT; 228 AA.

```

AC P38523;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update).
 DE GRPE protein homolog, mitochondrial precursor.
 GN GRPE1 OR GRPE OR YGE1 OR MEL1 OR YOR32W OR O5099.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 50-66; 91-96 AND 109-116.
 RX MEDLINE=94222047; PubMed=8168496;
 RA Bollinger L., Deloche O., Gluck B.S., Georgopoulos C., Jenoe P.,
 RA Kronidou N., Horst M., Morishima N., Schatz G.;
 RT "A mitochondrial homolog of bacterial GRPE interacts with
 RT mitochondrial hsp70 and is essential for viability.";
 RL EMO J. 13:1998-2006(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-X2180-1A;
 RX MEDLINE=94156045; PubMed=8112465;
 RA Ikeda E., Yoshida S., Mitsuizawa H., Uno I., Toh-E A.;
 RT "YGE1 is a yeast homologue of Escherichia coli grpe and is required
 RT for maintenance of mitochondrial functions.";
 RL FEBS Lett. 339:265-268(1994).
 RN [3]
 RP ERRATUM.
 RA Ikeda E., Yoshida S., Mitsuizawa H., Uno I., Toh-E A.;
 RL FEBS Lett. 343:181-181(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE=94294405; PubMed=8022808;
 RA Laloraya S., Gambill D.B., Craig E.A.;
 RT "A role for a eukaryotic GRPE-related protein, Mgelp, in protein
 RT translocation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6481-6485(1994).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / FY1679;
 RX MEDLINE=97127829; PubMed=8972580;
 RA Boyer J., Michaux G., Fairhead C., Gallon L., Dujon B.;
 RT "Sequence and analysis of a 26.9 kb fragment from chromosome XV of
 RT the yeast Saccharomyces cerevisiae.";
 RL Yeast 12:1575-1586(1996).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=95354663; PubMed=7628446;
 RA Westermann B., Prip-Buus C., Neupert W., Schwarz E.;
 RT "The role of the GRPE homologue, Mgelp, in mediating protein import
 RT and protein folding in mitochondria.";
 RL EMO J. 14:3452-3460(1995).
 CC -!- FUNCTION: COOPERATES WITH MITOCHONDRIAL HSP70 (SSC1) IN THE IMPORT
 CC OF PROTEINS FROM THE CYTOPLASM. SEEMS TO CONTROL THE NUCLEOTIDE-
 CC DEPENDENT BINDING OF SSC1 TO SUBSTRATE PROTEINS AND THE
 CC ASSOCIATION OF SSC1 WITH MIM4.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -!- PTM: THE N-TERMINUS IS BLOCKED.
 CC -!- SIMILARITY: BELONGS TO THE GRPE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X78350; CAA55145.1; -;
 DR EMBL: D26059; BAA05058.1; -;
 DR EMBL: U09565; AAA19253.1; -;
 DR EMBL: Z75140; CAA99452.1; -;

DR PIR: S43130; S43130.
 DR PIR: S41760; S41760.
 DR HSSP: P09372; 1DKG.
 DR SGD: S0005758; MGEL.
 DR InterPro: IPR000740; GRPE.
 DR Pfam: PF01025; GRPE; 1.
 DR PRINTS: PR00773; GREPROTEIN.
 DR PROSITE: PS01071; GRPE; 1.
 KW Chaperone; Mitochondrion; Transit peptide.
 FT TRANSIT 1
 FT CHAIN ? 228 MITOCHONDRION (POTENTIAL).
 FT SEQUENCE 228 AA; 26066 MW; 816A5E26C648E613 CRC64;
 SO
 Query Match 100.0%; Score 23; DB 1; Length 228;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KVGIV 5
 DB 219 KVGIV 223

Search completed: July 18, 2003, 16:02:18
 Job time : 6.52632 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:59:04 ; Search time 20 Seconds
(without alignments)
51.512 Million cell updates/sec

Title: US-10-054-968-6
Perfect score: 23
Sequence: 1 KVGIV 5

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protist:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rviro:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	100.0	81	17	08TR40
2	23	100.0	91	5	09U2A8
3	23	100.0	91	10	09AVW4
4	23	100.0	92	5	095W89
5	23	100.0	92	5	09VMU4
6	23	100.0	92	5	08TAK7
7	23	100.0	92	10	08RYU5
8	23	100.0	92	11	09CTW3
9	23	100.0	109	5	0967F1
10	23	100.0	120	5	0962X4
11	23	100.0	120	10	042093
12	23	100.0	122	10	042093
13	23	100.0	127	17	0974P3
14	23	100.0	152	16	09KT06
15	23	100.0	160	5	09BUE9
16	23	100.0			09bje9 plasmodium

17	23	100.0	162	17	09UX19	09ux19 sulfolobus
18	23	100.0	165	10	09SAW9	09saw9 aegilops ta
19	23	100.0	165	17	029242	029242 archaeoglob
20	23	100.0	167	16	08R614	08r614 thermotaneer
21	23	100.0	176	16	09ZMK0	09zmk0 helicobacte
22	23	100.0	177	2	08VNL8	08vnl8 enterococcu
23	23	100.0	177	16	08R7U4	08r7u4 thermotaneer
24	23	100.0	191	8	08WDD5	08wdd5 taenia soli
25	23	100.0	213	16	09ZM01	09zmq1 helicobacte
26	23	100.0	215	16	09KMZ6	09kzm6 vibrio chol
27	23	100.0	226	2	09RBP5	09rbp5 rhinococcus
28	23	100.0	231	16	08RG26	08rg26 fusobacteri
29	23	100.0	232	2	09RFP7	09rfp7 mycoplasma
30	23	100.0	239	16	0927J3	0927j3 listeria in
31	23	100.0	239	16	08YA29	08ya29 listeria mo
32	23	100.0	240	16	092MS3	092ms3 rhizobium m
33	23	100.0	240	16	08U6X5	08u6x5 agrobacteri
34	23	100.0	243	16	09KRX2	09kxx2 vibrio chol
35	23	100.0	243	16	08UK87	08uk87 agrobacteri
36	23	100.0	245	17	09UXS0	09uxs0 pyrococcus
37	23	100.0	245	17	08U311	08u311 pyrococcus
38	23	100.0	254	16	08X8P3	08x8p3 escherichia
39	23	100.0	259	5	09XTU1	09xtu1 caenorhabdi
40	23	100.0	259	17	0973D2	0973d2 sulfolobus
41	23	100.0	259	17	08TUC1	08tuc1 methanosarc
42	23	100.0	265	5	09VAX6	09vax6 drosophila
43	23	100.0	270	10	09FMU0	09fmu0 arabidopsis
44	23	100.0	272	5	09XW43	09xw43 caenorhabdi
45	23	100.0	276	16	P70970	P70970 bacillus su

ALIGNMENTS

RESULT 1

ID 08TR40 PRELIMINARY; PRT; 81 AA.
AC 08TR40:
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein MA1344.
GN MA1344.
OS Methanosarcina acetivorans.
OC Archaeae; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nisbaum C., Roy A., Endlitz M.G., MacDonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Alnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity."
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010804; AA004760.1; -
KW Hypothetical protein; Complete proteome.
SO SEQUENCE 81 AA; 8930 MW; 9F3E579587BB2707 CRC64;

Query Match 100.0%; Score 23; DB 17; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
 DB 32 KVGIV 36

RESULT 2

Q9U2A8 PRELIMINARY; PRT; 91 AA.
 AC Q9U2A8; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Y4B86A.2 protein.
 GN Y4B86A.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wall M.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;

RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology."
 RL Science.282:2012-2018(1998).
 DR EMBL: AL110490; CAB54440.1; -
 DR InterPro: IPR002674; Ribosomal_L37ae.
 DR Pfam: PF01780; Ribosomal_L37ae; 1.
 DR ProDom: PD006591; Ribosomal_L37ae; 1.
 DR TIGRFAMs: TIGR00280; L37a; 1.
 SQ SEQUENCE 91 AA; 10102 MW; 125510DECBE8E4C1 CRC64;

Query Match 100.0%; Score 23; DB 5; Length 91;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
 DB 7 KVGIV 11

RESULT 3

Q9AVM4 PRELIMINARY; PRT; 91 AA.
 AC Q9AVM4; 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 60S ribosomal protein L37A.
 GN RPL37A.
 OS Guillardia theta (Cryptomonas ph.).
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
 OX NCBI_Taxid=55529;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20087226; PubMed=10618395;
 RA Zanner S., Fraunholz M., Mastl J., Penny S.L., Beaton M.,
 RA Cavalier-Smith T., Maier U., Douglas S.;
 RT "Chloroplast protein and centrosomal genes, a tRNA intron, and odd
 RT telomeres in an unusually compact eukaryotic genome, the cryptomonad
 RT nucleomorph."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21223349; PubMed=11323671;
 RA Douglas S., Zanner S., Fraunholz M., Beaton M., Penny S., Deng L.,
 RA Wu X., Reith M., Cavalier-Smith T., Maier U.;
 RT "The highly reduced genome of an enslaved algal nucleus."
 RL Nature 410:1091-1096(2001).

DR EMBL: AJ010592; CAC27107.1; -
 DR InterPro: IPR002674; Ribosomal_L37ae.
 DR Pfam: PF01780; Ribosomal_L37ae; 1.
 DR ProDom: PD006591; Ribosomal_L37ae; 1.
 DR TIGRFAMs: TIGR00280; L37a; 1.
 KW Ribosomal protein.
 SQ SEQUENCE 91 AA; 10374 MW; 4803610F86C698D8 CRC64;

Query Match 100.0%; Score 23; DB 10; Length 91;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
 DB 7 KVGIV 11

RESULT 4

Q95W89 PRELIMINARY; PRT; 92 AA.
 AC Q95W89; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 60S ribosomal protein L37A.
 OS Branchiostoma belcheri (Amphioxus).
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 CC Branchiostoma.
 OX NCBI_Taxid=7741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chen Z., Zhang H.;
 RT "Cloning of 60S ribosomal protein L37A gene in amphioxus."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF363029; AA52799.2; -
 DR InterPro: IPR002674; Ribosomal_L37ae.
 DR Pfam: PF01780; Ribosomal_L37ae; 1.
 DR ProDom: PD006591; Ribosomal_L37ae; 1.
 DR TIGRFAMs: TIGR00280; L37a; 1.
 KW Ribosomal protein.
 SQ SEQUENCE 92 AA; 10366 MW; 87288ABCA1256D0F CRC64;

Query Match 100.0%; Score 23; DB 5; Length 92;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
 DB 7 KVGIV 11

RESULT 5

Q9VMU4 PRELIMINARY; PRT; 92 AA.
 AC Q9VMU4; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE CG5827 protein (RH41593p) (RE23595p).
 GN RPL37A OR CG5827.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayan A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Chertys J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Ibbagam C.,
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Wolley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiniker S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF0035608; AAF52217.1; -
 DR EMBL: AY011725; AAL49347.1; -
 DR EMBL: AY011189; AAL48811.1; -
 DR FLYbase: FBgn0028696; RPL37a.
 DR InterPro: IPR002674; Ribosomal_L37ae.
 DR Pfam: PF01780; Ribosomal_L37ae; 1.
 DR ProDom: PD006591; Ribosomal_L37ae; 1.
 DR TrifRAMS: TRIFR00280; L37a; 1.
 SO SEQUENCE 92 AA; 10311 MW; F8D85A34503C1FC6 CRC64;

Query Match Best Local Similarity 100.0%; Score 23; DB 5; Length 92;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KVGIV 5
 Db 7 KVGIV 11

RESULT 6
 0874K7 PRELIMINARY; PRT; 92 AA.
 AC 0874K7;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 OS L37a ribosomal protein.
 OC Taenia crassiceps.
 OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;

OC Cyclophyllidae; Taeniidae; Taenia.
 OX NCBI_TaxID=6207;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Goudal B., Coustau C., Milla G., Gabriel C.;
 RT "Complete coding sequence of Taenia crassiceps L37a ribosomal
 protein.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY081010; AAL83670.1; -
 KW RIBOSOMAL protein.
 SO SEQUENCE 92 AA; 10025 MW; F5E9AF15F3DC84B CRC64;

Query Match Best Local Similarity 100.0%; Score 23; DB 5; Length 92;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KVGIV 5
 Db 7 KVGIV 11

RESULT 7
 08RXU5 PRELIMINARY; PRT; 92 AA.
 AC 08RXU5;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 GN ATG60245.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Torimi M., Yu G., Bowser L.,
 RA Carinich P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamliya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda A., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Arabidopsis Full Length cDNA Clones";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY080669; AAL86345.1; -
 KW Hypothetical protein.
 SO SEQUENCE 92 AA; 10241 MW; D37A2D16631D410E CRC64;

Query Match Best Local Similarity 100.0%; Score 23; DB 10; Length 92;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KVGIV 5
 Db 7 KVGIV 11

RESULT 8
 09CYW3 PRELIMINARY; PRT; 92 AA.
 AC 09CYW3;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE 01-JUN-2001 (TREMblrel. 21, Last annotation update)
 GN RIBOSOMAL protein L37a.
 DE RPL37a.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinaawa A., Shibata K., Yoshino M., Itch M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi S., Fukuda S.,
 RA Atawa K., Itawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Futuro M., Kono H., Baldarelli R., Barsh G.,
 RA Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guinacich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK013246; BAB28742.1;
 DR MGI: 98068; Rpl37a.
 DR InterPro: IPR002674; Ribosomal_L37ae.
 DR Pfam: PF01780; Ribosomal_L37ae; 1.
 DR ProDom: PD006591; Ribosomal_L37ae; 1.
 DR TIGRFAMs: TIGR00280; L37a; 1.
 SQ SEQUENCE 92 AA; 10305 MW; 7139B53600E24CF6 CRC64;

Query Match 100.0%; Score 23; DB 11; Length 92;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
 |||||
 DB 7 KVGIV 11

RESULT 9
 090YTO PRELIMINARY; PRT; 92 AA.
 AC 090YTO;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Ribosomal protein L37a.
 OS Ictalurus punctatus (Channel catfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
 OC Ictaluridae; Ictalurus.
 OX NCBI_Taxid=7998;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Patterson A.P., Karsi A., Liu Z.J.;
 RT "Translational machinery of channel catfish: II. Complementary DNA and
 RT Expression of the Complete Set of 47 60S Ribosomal Proteins."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF401594; AAK95166.1;
 DR InterPro: IPR002674; Ribosomal_L37ae.
 DR Pfam: PF01780; Ribosomal_L37ae; 1.
 DR ProDom: PD006591; Ribosomal_L37ae; 1.
 DR TIGRFAMs: TIGR00280; L37a; 1.
 SQ SEQUENCE 92 AA; 10265 MW; 6E1FAF8600E24CF6 CRC64;

Query Match 100.0%; Score 23; DB 13; Length 92;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
 |||||
 DB 7 KVGIV 11

RESULT 10
 0967F1 PRELIMINARY; PRT; 109 AA.
 AC 0967F1;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein T27F7.3b.
 GN T27F7.3B.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium."
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Hallsworth K.;
 RT "The sequence of C. elegans cosmid T27F7."
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U58762; AAK39303.1;
 DR InterPro: IPR001950; TIF_SUI1.
 DR Pfam: PF01253; SUI1; 1.
 DR ProSite: PS50296; SUI1_2; 1.
 SQ SEQUENCE 109 AA; 12382 MW; 01A38C89159A4C79 CRC64;

Query Match 100.0%; Score 23; DB 5; Length 109;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
 |||||
 DB 95 KVGIV 99

RESULT 11
 0962X4 PRELIMINARY; PRT; 120 AA.
 AC 0962X4;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Multidrug resistance-associated protein (Fragment).
 GN MRP.
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_Taxid=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Luedeking A., Alpermann T., Koehler A.;
 RT "Differential expression of multidrug resistance related genes
 RT in Mytilus edulis."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF397142; AAK84397.1;
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00005; ABC_tran; 1.
 FT NON_TER 1 1
 FT NON_TER 120 120

SQL SEQUENCE 120 AA; 13406 MW; 2830DE6D0DB507F CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 5; Length 120;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
DB 67 KVGIV 71

RESULT 12

Q93YX9 PRELIMINARY; PRT; 120 AA.
AC Q93YX9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
OS Lipid transfer protein.
OC Davidia involucrata.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Cornales; Cornaceae; Davidioideae; Davidia.
OX NCBI_TaxID=16924;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Y., Chen L., Lin J., Li Y., Su Z., Chen F.;
RL Submitted (OCCT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY059472; AAL27855.1.
DR InterPro; IPR000528; Plant_LTP.
DR InterPro; IPR001768; TRY/amy1_inhbr.
DR Pfam; PF00234; try_alpha_amy1.1.
DR PROSITE; PS00597; PLANT_LTP; UNKNOWN.1.
SQ SEQUENCE 120 AA; 12057 MW; 7EA75BD099756DB CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 10; Length 120;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
DB 9 KVGIV 13

RESULT 13

Q42093 PRELIMINARY; PRT; 122 AA.
AC Q42093;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Multidrug resistance protein (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidops.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. COLUMBIA; TISSUE=SEEDLING;
RA Desprez T., Amselem J., Chapello H., Caboche M., Hofte H.;
RL Submitted (SEP-1993) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z26467; CAA81250.1.
FT NON_TER 1
FT NON_TER 122
SQ SEQUENCE 122 AA; 13262 MW; B2B9CD45274E5D06 CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 10; Length 122;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
DB 11111

DB 98 KVGIV 102

RESULT 14

Q974P3 PRELIMINARY; PRT; 127 AA.
AC Q974P3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein ST0616.
GN ST0616.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-JCM 10545 / 7;
RX PubMed-11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagil M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000983; BAB65614.1.
SQ Hypothetical protein; Complete proteome.
SQ SEQUENCE 127 AA; 14021 MW; 32ADB58C147EE86C CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 17; Length 127;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
DB 3 KVGIV 7

RESULT 15

Q9KT06 PRELIMINARY; PRT; 152 AA.
AC Q9KT06;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein VC1099.
GN VC1099.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umeyam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoli I., Sellers P.,
RA McDonald L., Uitterlinden T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004191; AAF94258.1.
DR TIGR; VC1099;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 152 AA; 17518 MW; A678A2643284D1D2 CRC64;

Query Match 100.0%; Score 23; DB 16; Length 152;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVGIV 5
 |||||
 Db 6 KVGIV 10

Search completed: July 18, 2003, 16:03:43
 job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:59:34 ; Search time 9.47368 Seconds
(Without alignments)
50.738 Million cell updates/sec

Title: US-10-054-968-6

Sequence: 1 KVGIV 5

Scoring table: BLASTSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR-73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	23	100.0	52	1	INGXA
2	23	100.0	52	2	S44469
3	23	100.0	52	2	S44470
4	23	100.0	84	2	S78266
5	23	100.0	91	2	T26996
6	23	100.0	91	2	C90116
7	23	100.0	92	1	R5R37
8	23	100.0	92	2	S42109
9	23	100.0	92	2	JN0875
10	23	100.0	92	2	S24170
11	23	100.0	93	2	S34661
12	23	100.0	152	2	F82240
13	23	100.0	156	2	E83844
14	23	100.0	156	2	G64447
15	23	100.0	162	2	A90210
16	23	100.0	165	1	D69377
17	23	100.0	176	2	F71958
18	23	100.0	213	2	C71965
19	23	100.0	215	2	C82493
20	23	100.0	220	2	D64316
21	23	100.0	228	2	S41760
22	23	100.0	239	2	AE1405
23	23	100.0	239	2	AE1781
24	23	100.0	240	2	H95873
25	23	100.0	240	2	G98155
26	23	100.0	240	2	AE3132
27	23	100.0	243	2	AG3188
28	23	100.0	243	2	C82393
29	23	100.0	245	2	A75056

30	23	100.0	248	2	C97652	hypothetical prote
31	23	100.0	254	2	S30742	hypothetical prote
32	23	100.0	254	2	D86068	hypothetical prote
33	23	100.0	254	2	A91222	hypothetical prote
34	23	100.0	269	2	B32410	mastocytoma protei
35	23	100.0	272	2	T26235	hypothetical prote
36	23	100.0	276	2	E69742	ABC transporter (A
37	23	100.0	277	2	T45265	ABC transporter (A
38	23	100.0	280	2	S16724	coenzyme F420 hydr
39	23	100.0	280	2	G69043	cobalt transporter A
40	23	100.0	288	2	AG1775	ABC transporter (A
41	23	100.0	288	2	AH1399	ABC transporter (A
42	23	100.0	289	1	SYEC5A	succinate-CoA liga
43	23	100.0	289	2	B90723	succinyl-CoA synth
44	23	100.0	289	2	A85574	succinyl-CoA synth
45	23	100.0	289	2	AG0591	succinyl-CoA synth

ALIGNMENTS

RESULT 1

INGXA

Insulin - alligator gar

C:Species: Lepisosteus spatula (alligator gar)

C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999

C:Accession: S09342

R:Pollock, H.G.; Kimmel, J.R.; Hamilton, J.W.; Rouse, J.B.; Ebner, K.E.; Lance, V.; R

Gen. Comp. Endocrinol. 67, 375-382, 1987

A:Title: Isolation and structures of alligator gar (Lepisosteus spatula) insulin and

A:Reference number: S07215; MUID:88030594; PMID:3311873

A:Accession: S09342

A:Molecule type: protein

A:Residues: 1-31;32-52 <POL>

C:Superfamily: Insulin

C:Keywords: hormone; pancreas

F:1-31/Domain: Insulin chain B #status experimental <BCH>

F:1-31;32-52/Product: Insulin A #status experimental <ACH>

F:7-38;19-51;37-42/Disulfide Bonds: #status predicted

Query Match 100.0%; Score 23; DB 1; Length 52;

Best Local Similarity 100.0%; Pred. No. 75;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5

Db 30 KVGIV 34

RESULT 2

S44469

Insulin II - North American paddlefish (Polyodon spathula)

C:Species: Polyodon spathula

C>Date: 27-Nov-1997 #sequence_revision 30-Jan-1998 #text_change 16-Jul-1999

C:Accession: S44469; S44467

R:Nguyen, T.M.; Mommson, T.P.; Mims, S.M.; Conlon, J.M.

Biochem. J. 300, 339-345, 1994

A:Title: Characterization of Insulins and proglucagon-derived peptides from a phyloge

A:Reference number: S44467; MUID:94271144; PMID:8002937

A:Accession: S44469

A:Molecule type: protein

A:Residues: 1-31 <NGU>

A:Experimental source: pancreas

A:Accession: S44467

A:Molecule type: protein

A:Residues: 32-52 <NGF>

A:Experimental source: pancreas

C:Superfamily: Insulin

C:Keywords: hormone; pancreas

F:1-31/Domain: Insulin II chain B #status predicted <BCH>

F:1-31;32-52/Product: Insulin II #status predicted <MAT>

F:32-52/Domain: Insulin II chain A #status predicted <ACH>

F:7-38,19-51,37-42/Disulfide bonds: #status predicted

Query Match 100.0%; Score 23; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
DB 30 KVGIV 34

RESULT 3

S44470
Insulin I2 - North American paddlefish (*Polyodon spathula*)

C:Species: *Polyodon spathula*

C:Date: 18-Sep-1997 #sequence_revision 01-May-1998 #text_change 16-Jul-1999

C:Accession: S44470; S44468

R:Nguyen, T.M.; Mommson, T.P.; Mims, S.M.; Conlon, J.M.

Biochem. J. 300, 339-345, 1994

A:Title: Characterization of insulins and proglucagon-derived peptides from a phylogenetic

A:Reference number: S44467; MUID:94271144; PMID:8002337

A:Accession: S44470

A:Molecule type: protein

A:Residues: 1-31 <NGU>

A:Experimental source: pancreas

A:Accession: S44468

A:Molecule type: protein

A:Residues: 32-52 <NGF>

A:Experimental source: pancreas

C:Superfamily: insulin

C:Keywords: hormone; pancreas

F:1-31/Domain: Insulin I2 chain B #status experimental <BCH>

F:1-31,32-52/Product: Insulin I2 #status experimental <MAT>

F:32-52/Domain: Insulin I2 chain A #status experimental <ACH>

F:7-38,19-51,37-42/Disulfide bonds: #status predicted

Query Match 100.0%; Score 23; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
DB 30 KVGIV 34

RESULT 4

S78266

ribosomal protein S17, chloroplast - *Odontella sinensis* chloroplast

C:Species: *Chloroplast Odontella sinensis*

C:Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 20-Jun-2000

C:Accession: S78266

R:Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.

Plant Mol. Biol. Rep. 13, 336-342, 1995

A:Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, *Odontella sinensis*

A:Reference number: S78238

A:Accession: S78266

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-84 <KOW>

A:Cross-references: EMBL:267753; NID:g1185127; PIDN:CA91639.1; PID:g1185156

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

C:Genetics:

A:Gene: rps17

A:Genome: chloroplast

C:Superfamily: *Escherichia coli* ribosomal protein S17

C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 100.0%; Score 23; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||

DB 6 KVGIV 10

RESULT 5

T26996

ribosomal protein L37a Y48B6a.2 [similarity] - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-May-2002

C:Accession: T26996

R:Wali, M.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20297

A:Accession: T26996

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-91 <RTL>

A:Cross-references: EMBL:AL110490; PIDN:CA54440.1; CESP:Y48B6a.2

A:Experimental source: clone Y48B6a

C:Genetics:

A:Gene: CESP:Y48B6a.2

C:Superfamily: rat ribosomal protein L37a

Query Match 100.0%; Score 23; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
DB 7 KVGIV 11

RESULT 6

C90116

60S ribosomal protein L37a [imported] - *Guillardia theta* nucleomorph

C:Species: nucleomorph *Guillardia theta*

A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 02-Aug-2002

C:Accession: C90116

R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.

Nature 410, 1091-1096, 2001

A:Title: The highly reduced genome of an enslaved algal nucleus.

A:Reference number: A99082; MUID:11323671; PMID:11323671

A:Accession: C90116

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-91 <DOU>

A:Cross-references: GB:A010592; NID:g12580789; PIDN:CAC27107.1; GSPDB:GN00151

C:Genetics:

A:Map position: 2

A:Genome: nucleomorph

C:Superfamily: rat ribosomal protein L37a

C:Keywords: nucleomorph

Query Match 100.0%; Score 23; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
DB 7 KVGIV 11

RESULT 7

R5R737

ribosomal protein L37a, cytosolic (validated) - rat

C:Species: *Rattus norvegicus* (Norway rat)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 21-Jul-2000

C:Accession: S05014; S11422

R:Tanaka, T.; Aoyama, Y.; Chan, Y.L.; Wool, I.G.

Eur. J. Biochem. 183, 15-18, 1989

A:Title: The primary structure of rat ribosomal protein L37a.

A:Reference number: S05014; MUID:89325328; PMID:2546769

A:Accession: S05014

A:Molecule type: mRNA
 A:Residues: 1-92 <TRAN>
 A:Cross-references: EMBL:X14069; NID:g57122; PIDN:CAA3232.1; PID:g57123
 A:Note: the protein is designated as ribosomal protein L37a
 R:Wittmann-Liebold, B.; Geisler, A.W.; Lin, A.; Wool, I.G.
 J. Supramol. Struct. 12, 425-433, 1979
 A:Title: Sequence of the amino-terminal region of rat liver ribosomal proteins S4, S6, S
 A:Reference number: S11413; MUID:80252792; PMID:398910
 A:Accession: S11422
 A:Molecule type: protein
 A:Residues: 2-29 <MTT>
 A:Note: the protein is designated as ribosomal protein L37a
 C:Superfamily: rat ribosomal protein L37a
 C:Keywords: protein biosynthesis; ribosome; zinc finger
 F:2-92/Product: ribosomal protein L37a #status experimental <MAT>

Query Match 100.0%; Score 23; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
 |||||
 DB 7 KVGIV 11

RESULT 8
 S42109
 ribosomal protein L37a, cytosolic - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999
 C:Accession: S42109; S34284
 R:Su, Y.; Raj, N.B.K.; Au, W.C.; Pilcha, P.M.
 Nucleic Acids Res. 21, 4400, 1993
 A:Title: Primary sequence of the mouse ribosomal protein L37a.
 A:Reference number: S42109; MUID:94021399; PMID:8415004
 A:Accession: S42109
 A:Molecule type: mRNA
 A:Residues: 1-92 <STV>
 A:Cross-references: EMBL:X73331; NID:g312413; PIDN:CAA51758.1; PID:g312414
 C:Superfamily: rat ribosomal protein L37a
 C:Keywords: protein biosynthesis; ribosome

Query Match 100.0%; Score 23; DB 2; Length 92;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
 |||||
 DB 7 KVGIV 11

RESULT 9
 JN0875
 ribosomal protein L37a - human
 C:Species: Homo sapiens (man)
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 13-Aug-1999
 C:Accession: JN0875; S26379
 R:Saba, D.P.; Tirumalai, P.S.; Scale, L.A.; Howells, R.D.
 Gene 132, 285-289, 1993
 A:Title: Human ribosomal protein L37a: cloning of the cDNA and analysis of differential
 A:Reference number: JN0875; MUID:94040775; PMID:8224875
 A:Accession: JN0875
 A:Molecule type: mRNA
 A:Residues: 1-92 <SAH>
 A:Cross-references: GB:L22154
 A:Experimental source: neuroblastoma
 R:Hoof, T.; Fislage, R.; Tuemmler, B.
 Nucleic Acids Res. 20, 5475, 1992
 A:Title: Primary sequence of the human ribosomal protein L37a.
 A:Reference number: S26379; MUID:93065220; PMID:1437567
 A:Accession: S26379
 A:Molecule type: preliminary
 A:Status: preliminary
 A:Molecule type: mRNA

A:Residues: 1-92 <HOO>
 A:Cross-references: EMBL:X66699; NID:g36133; PIDN:CAA47244.1; PID:g36134
 C:Superfamily: rat ribosomal protein L37a
 C:Keywords: ribosome; zinc finger
 F:39-60/Region: zinc finger

Query Match 100.0%; Score 23; DB 2; Length 92;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
 |||||
 DB 7 KVGIV 11

RESULT 10
 S24170
 ribosomal protein L37a - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 22-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
 C:Accession: S24170; S58475; S30354
 R:Toku, S.; Tanaka, T.
 Biochim. Biophys. Acta 1132, 88-90, 1992
 A:Title: The primary structure of chicken ribosomal protein L37a.
 A:Reference number: S24170; MUID:92379098; PMID:1380836
 A:Accession: S24170
 A:Molecule type: mRNA
 A:Residues: 1-92 <TOK>
 A:Note: Intron positions were determined by partial sequencing of genomic DNA
 R:Toku, S.; Tanaka, T.
 submitted to the EMBL Data Library, March 1992
 A:Reference number: S58475
 A:Accession: S58475
 A:Molecule type: mRNA
 A:Residues: 22-92 <TOW>
 A:Cross-references: EMBL:D10731; NID:g222861; PIDN:BA01575.1; PID:g441164
 R:Machida, M.; Toku, S.; Kennochi, N.; Tanaka, T.
 Eur. J. Biochem. 213, 77-80, 1993
 A:Title: The structure of the gene encoding chicken ribosomal protein L37a.
 A:Reference number: S30354; MUID:93238746; PMID:8477735
 A:Accession: S30354
 A:Molecule type: DNA
 A:Residues: 1-92 <MAC>
 A:Cross-references: EMBL:D14167; NID:g222865; PIDN:BA03209.1; PID:g222866
 A:Note: the authors translated the codon ATT for residue 10 as Lys, GTG for residue 1
 C:Genetics: 1/3; 44/3; 72/2
 C:Superfamily: rat ribosomal protein L37a
 C:Keywords: protein biosynthesis; ribosome

Query Match 100.0%; Score 23; DB 2; Length 92;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGIV 5
 |||||
 DB 7 KVGIV 11

RESULT 11
 S34661
 ribosomal protein L37a, cytosolic - turnip
 C:Species: Brassica rapa (turnip)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999
 C:Accession: S34661
 R:Song, S.; Choi, Y.
 submitted to the EMBL Data Library, July 1993
 A:Description: Nucleotide sequence of a B. rapa ribosomal protein L37a.
 A:Reference number: S34661
 A:Accession: S34661
 A:Molecule type: mRNA
 A:Residues: 1-93 <SON>
 A:Cross-references: EMBL:Z24739; NID:g395076; PIDN:CAA80864.1; PID:g395077

C:Superfamily: rat ribosomal protein L37a
C:Keywords: protein biosynthesis; ribosome

Query Match 100.0%; Score 23; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
Db 7 KVGIV 11

RESULT 12

F82240 conserved hypothetical protein VC1099 [Imported] - Vibrio cholerae (strain N16961 serog

C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: F82240

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: F82240
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-152 <HEI>
A:Cross-references: GB:AE004191; GB:AE003852; NID:99655568; PIDN:AA94258.1; GSPDB:GN001
C:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1099
A:Map position: 1
C:Superfamily: hypothetical protein H11205

Query Match 100.0%; Score 23; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
Db 6 KVGIV 10

RESULT 13

E83844 6,7-dimethyl-8-ribityllumazine synthase ribH [Imported] - Bacillus halodurans (strain C-

C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: E83844

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: E83844
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <STPO>
A:Cross-references: GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA05276.1; GSPDB:GN001
C:Experimental source: strain C-125
C:Genetics:
A:Gene: ribH
C:Superfamily: riboflavin synthase beta chain

Query Match 100.0%; Score 23; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
Db 15 KVGIV 19

RESULT 14

G64447

riboflavin synthase (EC 2.5.1.9) MJ1184 [similarity] - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Mar-2001

C:Accession: G64447

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Moese
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A:Reference number: A64300; MUID:96337939; PMID:8688087

A:Accession: G64447

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-156 <BUIL>

A:Cross-references: GB:U67559; GB:L77117; NID:g1591798; PIDN:AA99185.1; PID:g1591811

C:Genetics:

A:Map position: FOR1122695-1123165

A:Start codon: TTG

C:Superfamily: Methanobacterium thermoautotrophicum riboflavin synthase
C:Keywords: transferase

Query Match 100.0%; Score 23; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
Db 4 KVGIV 8

RESULT 15

A90210

hypothetical protein mets-like [Imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C:Accession: A90210

R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyaz, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.
Submitted to Genbank, April 2001

A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139

A:Accession: A90210

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-162 <KOR>

A:Cross-references: GB:AE006641; NID:g13813794; PIDN:AAK40936.1; GSPDB:GN00155

C:Genetics:

A:Gene: mets-like

Query Match 100.0%; Score 23; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGIV 5
|||||
Db 64 KVGIV 68

Search completed: July 18, 2003, 16:04:25
Job time: 10.4737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:56:19 ; Search time 28.4211 Seconds
(without alignments)
28.131 Million cell updates/sec

Title: US-10-054-968-7

Perfect score: 36

Sequence: 1 FGPKNF 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	6	AAW44380	Vancomycin resist
2	36	100.0	180	AAW44388	Vancomycin resist
3	36	100.0	273	AAU34054	Staphylococcus aur
4	36	100.0	286	AAU36641	Staphylococcus aur
5	36	100.0	288	ABP38447	Staphylococcus epi
6	36	100.0	289	AAU34952	Enterococcus faeca
7	33	91.7	84	AAU01684	Human polypeptide
8	32	88.9	279	AAU81709	Streptococcus pneu
9	32	88.9	279	AAU37847	Streptococcus pneu
10	32	88.9	279	AAU01016	CPE 15 protein seq

11	32	88.9	280	23	ABP26997
12	32	88.9	280	23	ABP26998
13	32	88.9	288	23	ABP35584
14	31	86.1	1014	22	ABP36655
15	31	86.1	1015	19	ABP3675
16	31	86.1	1015	22	ABP3675
17	31	86.1	1043	22	ABP3664
18	30	83.3	14	22	AAU98670
19	30	83.3	14	22	AAU00631
20	30	83.3	24	18	AAU30912
21	30	83.3	24	21	AAU36384
22	30	83.3	24	22	AAU05558
23	30	83.3	42	23	ABP26747
24	30	83.3	52	22	AAU6547
25	30	83.3	71	22	AAU006514
26	30	83.3	77	22	ABP50277
27	30	83.3	95	22	AAU66608
28	30	83.3	100	22	AAU11207
29	30	83.3	106	21	AAU59024
30	30	83.3	106	23	ABP41091
31	30	83.3	109	22	AAU75741
32	30	83.3	115	22	AAU13439
33	30	83.3	120	22	AAU11391
34	30	83.3	129	22	AAU09751
35	30	83.3	140	22	AAU01652
36	30	83.3	148	22	AAU6490
37	30	83.3	161	21	AAU52229
38	30	83.3	193	18	AAU30905
39	30	83.3	193	18	AAU30906
40	30	83.3	193	18	AAU30907
41	30	83.3	193	13	AAU58395
42	30	83.3	193	21	AAU58370
43	30	83.3	193	21	AAU58372
44	30	83.3	193	21	AAU58373
45	30	83.3	193	21	AAU58374

ALIGNMENTS

RESULT 1	
AAW44380	standard; peptide: 6 AA.
AAW44380:	
10-JUN-1998	(first entry)
XX	Vancomycin resistant Enterococcus faecium specific epitope SEQ ID NO.7.
XX	Vancomycin resistant Enterococcus faecium; VRE; epitope; immunogen;
XX	methicillin resistant Staphylococcus aureus; MRSA; transporter protein.
OS	Enterococcus faecium.
XX	
PN	WO9801154-A2.
XX	
PD	15-JAN-1998.
XX	
PF	07-JUL-1997; 97WO-GB01830.
XX	
PR	06-JUL-1996; 96GB-0014274.
XX	
PA	(NEUT-) NEUTEC PHARM PLC.
XX	
PI	Burnie JP, Matthews RC;
XX	
DR	WPI; 1998-100821/09.
XX	
PT	Treating and diagnosing bacterial and fungal infection with ABC
PT	transporter protein - Or neutralising or binding agents, and new
PT	Staphylococcal proteins, particularly for infections caused by drug
PT	resistant Staphylococci and Enterococci

XX Claim 9; Page 38; 62pp; English.

PS The present sequence represents a vancomycin resistant Enterococcus

XX faecium (VRE) specific epitope. The present invention describes the

CC therapeutic or diagnostic use of a bacterial or fungal ABC transporter

CC protein, or its immunogenic fragments, in human or veterinary medicine.

CC Binding agents, specific for an ABC transporter protein, are used in

CC standard binding assays to detect the ABC transporter protein, while

CC ABC transporter proteins are used similarly to detect specific

CC antibodies in patient sera, specifically for diagnosis of Staphylococci

CC and Enterococci infections. Neutralising agents specific for ABC

CC transporter proteins, specifically antibodies (Ab), are used to treat

CC such infections (optionally when coupled to a drug) while ABC

CC transporter proteins are used as immunogens to protect against

CC infection. Ab raised against ABC transporter proteins are effective

CC against multiple drug resistant strains of bacteria (specifically

CC vancomycin and/or methicillin resistant strains) for which no drug

CC therapy is available.

CC

XX

SQ Sequence 6 AA;

Query Match 100.0%; Score 36; DB 19; Length 6;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6

DB 1 FGPKNF 6

RESULT 2

AAW44388

ID AAW44388 standard; Protein; 180 AA.

XX

AC AAW44388;

XX

DT 10-JUN-1998 (first entry)

XX

DE Vancomycin resistant Enterococcus faecium expression product.

XX

KM Vancomycin resistant Enterococcus faecium; VRE; epitope; immunogen;

XX methicillin resistant Staphylococcus aureus; MRSA; transporter protein.

XX

OS Enterococcus faecium.

XX

PN WO9801154-A2.

XX

PD 15-JAN-1998.

XX

PF 07-JUL-1997; 97WO-GB01830.

XX

PR 06-JUL-1996; 96GB-0014274.

XX

PA (NEUT-) NEUTEC PHARM PLC.

XX

PI Burnie JP, Matthews RC;

XX

DR WPI: 1998-100821/09.

XX

DR N-PDB; AAV15410.

XX

PT Treating and diagnosing bacterial and fungal infection with ABC

XX transporter protein - or neutralising or binding agents, and new

PT Staphylococcal proteins, particularly for infections caused by drug

XX resistant Staphylococci and Enterococci

XX

PS Example; Page 36; 62pp; English.

XX

CC The present sequence represents a vancomycin resistant Enterococcus

CC faecium (VRE) expression product. The present invention describes the

CC therapeutic or diagnostic use of a bacterial or fungal ABC transporter

CC protein, or its immunogenic fragments, in human or veterinary medicine.

CC Binding agents, specific for an ABC transporter protein, are used in

CC standard binding assays to detect the ABC transporter protein, while

CC ABC transporter proteins are used similarly to detect specific

CC antibodies in patient sera, specifically for diagnosis of Staphylococci

CC and Enterococci infections. Neutralising agents specific for ABC

CC transporter proteins, specifically antibodies (Ab), are used to treat

CC such infections (optionally when coupled to a drug) while ABC

CC transporter proteins are used as immunogens to protect against

CC infection. Ab raised against ABC transporter proteins are effective

CC against multiple drug resistant strains of bacteria (specifically

CC vancomycin and/or methicillin resistant strains) for which no drug

CC therapy is available.

CC

XX

SQ Sequence 180 AA;

Query Match 100.0%; Score 36; DB 19; Length 180;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6

DB 113 FGPKNF 118

RESULT 3

AAU34054

ID AAU34054 standard; Protein; 273 AA.

XX

AC AAU34054;

XX

DT 14-FEB-2002 (first entry)

XX

DE Staphylococcus aureus cellular proliferation protein #330.

XX

KM Antisense: prokaryotic cellular proliferation protein;

XX antibiotic; antibacterial; drug design.

XX

OS Staphylococcus aureus.

XX

PN WO200170955-A2.

XX

PD 27-SEP-2001.

XX

PF 21-MAR-2001; 2001WO-US09180.

XX

PR 21-MAR-2000; 2000US-191078P.

XX

PR 23-MAY-2000; 2000US-206848P.

XX

PR 26-MAY-2000; 2000US-207727P.

XX

PR 23-OCT-2000; 2000US-242578P.

XX

PR 27-NOV-2000; 2000US-253625P.

XX

PR 22-DEC-2000; 2000US-257931P.

XX

PR 16-FEB-2001; 2001US-269308P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX

DR WPI: 2001-611495/70.

XX

DR N-PDB; AAS51913.

XX

PT New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids -

XX

PS Example 3; Seq ID No 5550; 51pp; English.

XX

CC The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 273 AA;

Query Match 100.0%; Score 36; DB 22; Length 273;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FGPKNF 6
 |||||

DB 94 FGPKNF 99

RESULT 4

AAU36641 standard; Protein: 286 AA.

AC AUA36641;

DT 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #811.

KW Antisense: prokaryotic cellular proliferation protein;

KW antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

DR WPI: 2001-611495/70.

DR N-PSDB: AAS54500.

PS New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 12234; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*

CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 286 AA;

Query Match 100.0%; Score 36; DB 22; Length 286;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FGPKNF 6
 |||||

DB 109 FGPKNF 114

RESULT 5

ABP38447 standard; Protein: 288 AA.

AC ABP38447;

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3292.

KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;

KW antibacterial; gene therapy.

OS Staphylococcus epidermidis.

PN US6380370-B1.

PD 30-APR-2002.

PF 13-AUG-1998; 98US-0134001.

PR 14-AUG-1997; 97US-055779P.

PR 08-NOV-1997; 97US-064964P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Doucette-Stamm LA, Bush D;

PI WPI: 2002-381255/41.

DR N-PSDB: ABN90992.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis

PT polypeptide, useful for diagnosing and treating bacterial infections -

XX Disclosure: SEQ ID 3292; 267pp; English.

PS ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading

CC frame (ORF) nucleic acid sequences which encode the amino acid sequences

CC given in ABP35124 to ABP37960. The S. epidermidis sequences have

CC antibacterial activity and can be used in gene therapy. The sequences

CC can also be used in the diagnosis and treatment of bacterial infections,

CC particularly S. epidermidis infections. The sequences can be used to

CC screen for compounds able to interfere with the S. epidermidis life

CC cycle or inhibit S. epidermidis infection.

CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the

CC USPTO web site.

XX Sequence 288 AA;

XX

Query Match 100.0%; Score 36; DB 23; Length 288;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 |||||
 Db 111 FGPKNF 116

RESULT 6

AAU34952
 ID AAU34952 standard; Protein: 289 AA.

AC AAU34952;

DT 14-FEB-2002 (first entry)

DE Enterococcus faecalis cellular proliferation protein #239.

KW Antisense; prokaryotic cellular proliferation protein;

KM antibiotic; antibacterial; drug design.

OS Enterococcus faecalis.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

DR WPI; 2001-611495/70.

DR N-PSDB; AAS52811.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

PS Example 3; Seq ID No 10545; 51pp; English.

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,

CC and to obtain antibodies capable of binding to the expressed proteins.

CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen

CC for homologous nucleic acids which are required for cell proliferation in

CC a wide variety of organisms. The present sequence represents an

CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 289 AA;

Query Match 100.0%; Score 36; DB 22; Length 289;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 |||||
 Db 109 FGPKNF 114

RESULT 7

AAO01684
 ID AAO01684 standard; Protein: 84 AA.

AC AAO01684;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 15576.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

OS nervous system disorders; arthritis; inflammation.

PN Homo sapiens.

PD WO200164835-A2.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PR (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

PI WPI; 2001-514838/56.

DR N-PSDB; AAI61615.

XX Isolated nucleic acids and polypeptides, useful for preventing

PT diagnosing and treating e.g. leukaemia, inflammation and immune

PS disorders -

XX Claim 20; SEQ ID NO 15576; 139pp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AA179941-AA193841) and

CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 84 AA;

CC Query Match 91.7%; Score 33; DB 22; Length 84;

CC Best Local Similarity 83.3%; Pred. No. 23;

CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CC OY 1 FGPKNF 6

CC |||||

CC Db 3 FGPKNF 8

RESULT 8
 ID AAY81709 standard; Protein: 279 AA.
 XX
 AC AAY81709;
 XX
 DT 02-JUN-2000 (first entry)
 XX
 DE Streptococcus pneumoniae protein sequence ID2.
 XX
 KW Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;
 KM bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;
 KW kidney disease; diabetes; immunosuppressive disorder; otitis media;
 KM pneumococcal septicemia; sinusitis; meningitis; therapy.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200006738-A2.
 XX
 PD 10-FEB-2000.
 XX
 PE 27-JUL-1999; 99WO-GB02452.
 XX
 PR 27-JUL-1998; 98GB-0016336.
 PR 19-MAR-1999; 99US-0125329.
 XX
 PA (MICR-) MICROBIAL TECHNIQS LTD.
 XX
 PI Le Page RWF, Wells JM, Hanniffy SB, Hansbro PM;
 XX
 DR WPI: 2000-195301/17.
 DR N-PSDB; AA291805.
 XX
 PT Streptococcal proteins and polynucleotides useful for diagnosis,
 PT treatment and prophylaxis of bacterial infections
 XX
 PS Claim 2; Page 40; 76pp; English.
 XX
 CC This sequence represents a Streptococcus pneumoniae protein of the
 CC invention. The proteins (or their homologues, derivatives and/or
 CC fragments) are useful as immunogens or antigens. Immunogenic or antigenic
 CC compositions comprising the proteins are useful as vaccines and also in
 CC diagnostic assays. The sequences are useful for the detection or
 CC diagnosis of S. pneumoniae infection, by contacting a sample to be tested
 CC with them. Agents capable of antagonising, inhibiting or interfering with
 CC the function or expression of the protein or polypeptide are useful in
 CC medical compositions in the treatment or prophylaxis of S. pneumoniae
 CC infection. As the sequences can be used to treat S. pneumoniae infection,
 CC young children, the elderly, and in patients with predisposing conditions
 CC such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,
 CC or with immunosuppressive disorders, especially AIDS. They can also be
 CC used to treat pneumococcal septicemia, otitis media, sinusitis, and
 CC meningitis.
 CC
 SO Sequence 279 AA;
 XX
 OY 1 FGPKNF 6
 DB 109 FGPONF 114
 XX
 RESULT 9
 ID AAU37847 standard; Protein: 279 AA.
 XX
 AC AAU37847;
 XX

DT 14-FEB-2002 (first entry)
 XX
 DE Streptococcus pneumoniae cellular proliferation protein #276.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KM antibiotic; antibacterial; drug design.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 DE 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zysek JM, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI: 2001-611495/70.
 DR N-PSDB; AAS55706.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 13440; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 SO Sequence 279 AA;
 XX
 OY 1 FGPKNF 6
 DB 109 FGPONF 114
 XX
 RESULT 10
 ID AAM01016 standard; Protein: 279 AA.
 XX
 AC AAM01016;
 XX

DT 02-OCT-2001 (first entry)
XX CFE 15 protein sequence.
XX Antibacterial; vaccine; gene therapy; bacterial cell wall viability;
KM CFE; CEG; Conserved Essential Gene; bacterial infection;
KM antisense therapy; antibiotic resistance.
XX Streptococcus pneumoniae.
OS
XX
PN WO200149721-A2.
XX
XX 12-JUL-2001.
XX
XX 29-DEC-2000; 2000WO-0535604.
XX
XX 30-DEC-1999; 99US-0174089.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Dougherty TJ, Pucci MJ, Dougherty BA, Davison DB, Bruccoleri RE;
PI Thanassi JA;
XX
XX WPI: 2001-496721/54.
DR N-PSDB; AAH90715.
XX
XX Nucleic acids encoding conserved essential genes involved in bacterial
PT replication which are potential targets for the treatment of antibiotic
PT resistant bacterial infections -
XX
PS Claim 27; Pages 254-255; 380pp; English.
XX
XX The present invention relates to nucleic acids (AAH90701-AAH90918)
CC encoding polypeptides (AAH01002-AAH0114), which are essential for the
CC viability of a bacterial cell wall. The acronym CFE stands for "Conserved
CC Expression", where CEG stands for "Conserved Essential Gene". The nucleic
CC acids are useful for detecting the presence of proteins essential for the
CC viability of a bacterial cell wall in samples such as cells, tissues,
CC biological fluids, blood, serum, nose, ear or throat swabs with ligands,
CC and for detecting corresponding target nucleic acid molecules with
CC complementary sequences. The nucleic acids are also useful for
CC determining whether a genomic nucleotide sequence of interest is
CC essential for viability of a bacterial cell or whether it resides within
CC an operon, by integrating an exogenous nucleotide sequence comprising a
CC portion of an open reading frame of the genomic sequence of interest
CC (comprising 200-500 base pairs) into the genomic sequence of interest
CC which confers a selectable phenotype to the cell, and determining cell
CC viability with a selection agent such as chloramphenicol. The nucleic
CC acids and proteins are also useful as vaccines and for treating bacterial
CC infections with gene therapy and antisense therapy. The nucleic acids
CC also enable identification of targets suitable for the treatment of
CC antibiotic resistant bacterial infections.
CC
XX
SO Sequence 279 AA;
OY
DB . 109 FGPQNF 114
OY 1 FGPKNF 6
DB . 109 FGPQNF 114
Query Match 88.9%; Score 32; DB 22; Length 279;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
RESULT 11
ABP26997
ID ABP26997 standard; Protein; 280 AA.
XX
XX ABP26997;
AC
XX
XX 02-JUL-2002 (first entry)
DT
XX Streptococcus polypeptide SEQ ID NO 3170.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KM group A streptococcus; Streptococcus pyogenes; antibacterial;
KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
XX Streptococcus agalactiae.
XX
PN WO200234771-A2.
XX
XX 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB04789.
XX
XX 27-OCT-2000; 2000GB-0026333.
XX
XX 24-NOV-2000; 2000GB-0028727.
XX
XX 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
XX Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettein H;
XX
XX WPI: 2002-352536/38.
DR N-PSDB; ABN67628.
XX
XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 1; Page 3470; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (1), nucleic acids encoding (1), ABN6604-ABN71526 and
CC antibodies that bind (1) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (1) are used to detect Streptococcus in a
CC biological sample. (1) is used to determine whether a compound binds to
CC (1). A composition comprising (1) or a nucleic acid encoding (1) may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (1) may be used to recombinantly produce (1) and may be
CC used in gene therapy. Antibodies to (1) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
CC
XX
SO Sequence 280 AA;
OY
DB . 109 FGPQNF 114
OY 1 FGPKNF 6
DB . 109 FGPQNF 114
Query Match 88.9%; Score 32; DB 23; Length 280;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
RESULT 12
ABP26998
ID ABP26998 standard; Protein; 280 AA.
XX
XX ABP26998;
AC
XX
XX 02-JUL-2002 (first entry)
DT
XX Streptococcus polypeptide SEQ ID NO 3172.
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KM group A streptococcus; Streptococcus pyogenes; antibacterial;

KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus pyogenes.
 OS
 PN WO200234771-A2.
 XX
 XX 02-MAY-2002.
 PD
 XX
 PF 29-OCT-2001; 2001WO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Telford J, Maignan V, Margart Ros VI, Grandi G, Fraser C;
 PI Tetelin H;
 XX
 DR MPI: 2002-352536/38.
 DR N-PSDB; ABN67629.
 XX
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1; Page 3470; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 XX
 SQ Sequence 280 AA:
 Query Match 88.9%; Score 32; DB 23; Length 280;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 FGPKNF 6
 III:II
 Db 109 FGPKNF 114
 RESULT 13
 ABB53584
 ID ABB53584 standard; Protein; 288 AA.
 XX
 AC ABB53584;
 XX
 DT 16-MAY-2002 (first entry)
 XX
 DE Lactococcus lactis protein yche.
 XX
 KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX
 OS Lactococcus lactis IL1403.
 XX
 PN FR2807446-A1.

XX
 PD 12-OCT-2001.
 XX
 PF 11-APR-2000; 2000FR-0004630.
 XX
 PR 11-APR-2000; 2000FR-0004630.
 XX
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX
 PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX
 DR MPI: 2002-043418/06.
 XX
 PT New nucleotide sequence useful in the identification or Lactococcus
 PT lactis and related species -
 XX
 PS Claim 6; SEQ ID No 286; 2504pp; French.
 XX
 CC The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 288 AA:
 Query Match 88.9%; Score 32; DB 23; Length 288;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 FGPKNF 6
 III:II
 Db 108 FGPKNF 113

RESULT 14
 AAB73665
 ID AAB73665 standard; Protein; 1014 AA.
 XX
 AC AAB73665;
 XX
 DT 11-SEP-2001 (first entry)
 XX
 DE Mouse ageing inhibitory protein #2, SEQ ID NO:2.
 XX
 KW Mouse; ageing inhibitory protein; recombinant production;
 KW transgenic animal; detection; drug screening; ligand identification;
 KW ageing-related disorder; premature ageing.
 XX
 OS Mus sp.
 XX
 PN WO200138529-A1.
 XX
 PD 31-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-JP08121.
 XX
 PR 19-NOV-1999; 99JP-0329649.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Shiraiishi N, Sekine S, Nabeshima Y, Fujimori T, Ito S;
 XX
 DR MPI: 2001-367689/38.
 DR N-PSDB; AAH24194.
 XX
 PT Polypeptide for treatment of premature aging -

XX Claim 1; Page 62-66; 94pp; Japanese.
 PS
 CC The invention relates to 3 novel mouse proteins which inhibit ageing
 CC (AAB73664-AAB73666) and to the nucleic acids encoding them (AAH24193-
 CC AAH24195). The invention also relates to recombinant vectors, host
 CC cells and transgenic non-human animals comprising DNA encoding a
 CC mouse ageing inhibitory protein of the invention; the recombinant
 CC production of a mouse ageing inhibitory protein; antibodies against
 CC the ageing inhibitory proteins; methods for immunologically assaying
 CC the proteins; and methods for identifying ligands or other modulators
 CC of the ageing inhibitory proteins; and the ligands and modulators thus
 CC identified. The ageing inhibitory proteins, nucleic acids encoding
 CC them, and antibodies, ligands and modulators, may be used in the
 CC diagnosis and treatment of ageing (particularly premature ageing),
 CC ageing-related disorders. The present sequence represents a
 CC 1014 amino acid mouse ageing inhibitory protein (42).
 CC
 SQ Sequence 1014 AA:
 Query Match 86.1%; Score 31; DB 22; Length 1014;
 Best Local Similarity 83.3%; Pred. No. 6.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 FGPKNF 6
 ||| ||
 Db 343 FGPKNF 348
 RESULT 15
 AAM63675
 ID AAM63675 standard; Protein: 1015 AA.
 XX
 AC AAM63675;
 XX
 DT 17-NOV-1998 (first entry)
 XX
 DE Polypeptide having ageing-retarding activity.
 XX
 KW Ageing-retarding activity; ligand; receptor; screening; inhibitor;
 KW augmenting; premature ageing; Down's syndrome; Turner's syndrome;
 KW Werner's disease; Rosamond-Thomson disease; gene therapy; livestock;
 KW domestic animal; human.
 XX
 OS Homo sapiens.
 XX
 PN WO9829544-A1.
 XX
 PD 09-JUL-1998.
 XX
 PF 12-DEC-1997; 97WO-JF04585.
 XX
 PR 31-JUL-1997; 97JP-0205815.
 PR 26-DEC-1996; 96JP-0347871.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Iida A, Kuroo M, Nabeshima Y, Sekine S;
 XX
 DR WPI; 1998-388125/33.
 DR N-PDB; AAV39131.
 XX
 PT Polypeptide which retards ageing - for diagnosis and treatment of
 PT ageing-related diseases and for improvement of livestock breeds.
 XX
 PS Claim 2; Pages 132-137; 225pp; Japanese.
 XX
 CC This represents a polypeptide that has an ageing-retarding activity.
 CC The invention provides nucleic acid sequences (AAV39127 to AAV39131)
 CC encoding polypeptides (AAM63671 to AAM63675) with ageing-retarding
 CC activity. Expression vectors comprising the nucleic acid sequences can be
 CC used to transform host cells for the recombinant production of such
 CC polypeptides. The peptides and their fragments may be used for

CC investigating possible ligands and receptors which bind to them, and for
 CC screening potential inhibitors of binding of the peptides to these
 CC ligands and receptors. They can also be used for screening potential
 CC augmenters and modifiers of expression of the genes coding for the
 CC peptides. They are used in the treatment and prevention of diseases
 CC involving premature ageing, such as Down's syndrome, Turner's syndrome,
 CC Werner's disease, or Rosamond-Thomson disease, and of other diseases of
 CC adults, by administering the peptides or of binding inhibitors or gene
 CC expression modifiers identified by screening as above or by gene therapy.
 CC Gene therapy may also be used for producing improved stock of domestic
 CC animals and livestock.
 CC
 SQ Sequence 1015 AA:
 Query Match 86.1%; Score 31; DB 19; Length 1015;
 Best Local Similarity 83.3%; Pred. No. 6.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 FGPKNF 6
 ||| ||
 Db 345 FGPKNF 350
 Search completed: July 18, 2003, 16:01:51
 Job time : 29.4211 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:59:34 ; Search time 11.3684 Seconds
(Without alignments)
50.738 Million cell updates/sec

Title: US-10-054-968-7

Perfect score: 36

Sequence: 1 FGPKNF 6

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

283224

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	266	2 H90018	hypothetical prote
2	33	91.7	461	2 H98255	DNA modification m
3	33	91.7	461	2 AC3029	DNA methyltransfer
4	32	88.9	279	2 C95259	ABC transporter, A
5	32	88.9	279	2 F98124	hypothetical prote
6	32	88.9	288	2 F86559	ABC transporter AT
7	32	88.9	912	2 T02892	hypothetical prote
8	31	86.1	307	2 F84502	En/Spm transposon
9	31	86.1	533	2 F86441	probable cytochrom
10	30	83.3	77	1 GYR1	cysteine-rich inte
11	30	83.3	77	2 JC2431	cysteine-rich prot
12	30	83.3	77	2 G02666	cysteine-rich prot
13	30	83.3	147	2 S67982	hemoglobin beta-II
14	30	83.3	192	2 S38879	LTM-domain protein
15	30	83.3	192	2 A49648	cysteine-rich prot
16	30	83.3	193	2 S53580	cysteine-rich prot
17	30	83.3	193	2 S12658	cysteine-rich prot
18	30	83.3	194	2 S41761	cysteine-rich prot
19	30	83.3	194	2 S52335	beta-cysteine-rich
20	30	83.3	208	2 T33341	hypothetical prote
21	30	83.3	213	2 E96943	similar to hypothe
22	30	83.3	215	2 F72412	hypothetical prote
23	30	83.3	260	2 B71234	probable transport
24	30	83.3	270	2 T06992	translation initia
25	30	83.3	276	2 E69742	ABC transporter (A
26	30	83.3	288	2 T22789	hypothetical prote
27	30	83.3	288	2 AG1775	ABC transporter (A
28	30	83.3	288	2 AH1399	ABC transporter (A
29	30	83.3	301	2 S50737	probable membrane

30	30	83.3	307	2 T17104	translation initia
31	30	83.3	309	2 S67785	probable membrane
32	30	83.3	325	2 B89909	GMP reductase (EC
33	30	83.3	330	2 A70422	hypothetical prote
34	30	83.3	331	2 D96515	hypothetical prote
35	30	83.3	341	2 AE1079	ornithine carbamoy
36	30	83.3	343	2 T42528	hypothetical prote
37	30	83.3	347	2 T42528	BDL protein homol
38	30	83.3	347	2 T38998	probable galactosy
39	30	83.3	349	2 AD3650	trans-1,2-dihydrob
40	30	83.3	356	2 H72803	integrase protein
41	30	83.3	369	2 AB1537	hypothetical prote
42	30	83.3	473	2 C86841	hypothetical prote
43	30	83.3	504	2 G84251	hypothetical prote
44	30	83.3	579	2 A70954	hypothetical prote
45	30	83.3	693	2 S46417	heat shock protein

ALIGNMENTS

RESULT 1
H90018
hypothetical protein SA2020 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: H90018
R:Kuroda, M.; Ohba, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsuka, K.
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: H90018
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <KUR>
A:Cross-references: GB:BA000018; PID:g13702021; PIDN:BA843313.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA2020
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 100.0%; Score 36; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. NO. 4.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FGPKNF 6
Db 109 FGPKNF 114

RESULT 2
H98255
DNA modification methyltransferase mxdAI (AF051092) [imported] - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: H98255
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: H98255
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-461 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK89570.1; PID:g15159455; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_1997
A:Map position: linear chromosome

Query Match 91.7%; Score 33; DB 2; Length 461;

Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
111:11
Db 39 FGPRNF 44

RESULT 3

DNA methyltransferase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AC3029
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McCell
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AC3029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-467 <KUR>
A:Cross-references: GB:AE008689; PIDN:AL44649.1; PID:917742274; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu3839
A:Map position: linear chromosome

Query Match 91.7%; Score 33; DB 2; Length 467;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
111:11
Db 45 FGPRNF 50

RESULT 4

C95259
ABC transporter, ATP-binding protein SP2220 [imported] - Streptococcus pneumoniae (strai
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 17-May-2002
C:Accession: C95259
R:Teitelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a Virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: C95259
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-279 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK76268.1; PID:914973730; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2220
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Query Match 88.9%; Score 32; DB 2; Length 279;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
111:11
Db 109 FGPRNF 114

RESULT 5
F98124
hypothetical protein ABC-NBP [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 17-May-2002
C:Accession: F98124
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAnren, S.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: F98124
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-279 <KUR>
A:Cross-references: GB:AE007317; PIDN:AL00827.1; PID:915459731; GSPDB:GN00174
C:Genetics:
A:Gene: ABC-NBP
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Query Match 88.9%; Score 32; DB 2; Length 279;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
111:11
Db 109 FGPRNF 114

RESULT 6

F86659
ABC transporter ATP-binding protein yche [imported] - Lactococcus lactis subsp. lacti
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: F86659
R:Boletun, A.; Winkler, P.; Mauger, S.; Jallou, O.; Malarne, K.; Weissenbach, J.; Eh
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A:Reference number: A86625; MUID:21235186; PMID:11537471
A:Accession: F86659
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-288 <STO>
A:Cross-references: GB:AE005176; PID:912723140; PIDN:AAK04376.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yche
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Query Match 88.9%; Score 32; DB 2; Length 288;
Best Local Similarity 83.3%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
111:11
Db 108 FGPRNF 113

RESULT 7

T02892
hypothetical protein T13J8.20 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
C:Accession: T02892
R:Bevan, M.; Pohl, T.; Welzenegger, T.; Hobeisel, J.; Meves, H.W.; Mayer, K.F.X.; Sch
submitted to the Protein Sequence Database, February 1999
A:Reference number: T02892
A:Accession: T02892
A:Molecule type: DNA
A:Residues: 1-912 <BEV>
A:Cross-references: EMBL:AL035524

Query Match 88.9%; Score 32; DB 2; Length 279;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

A:Experimental source: cultivar Columbia; BAC clone T13J8
C:Genetics:
A:Map position: 4
A:Introns: 260/3; 279/2; 357/3; 400/3; 423/3; 510/2; 555/3; 619/1; 667/3; 702/3; 724/2;
A:Note: T13J8.20

Query Match 88.9%; Score 32; DB 2; Length 912;
Best Local Similarity 83.3%; Pred. No. 91;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
Db 200 FGPKNF 205

RESULT 8

En/Spm transposon protein (mosaic-like protein) [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84502
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.;
Euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617137
A:Accession: F84502
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-307 <STO>
A:Cross-references: GB:AE002093; NID:94038061; PIDN:AAC97242.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g12300
A:Map position: 2

Query Match 86.1%; Score 31; DB 2; Length 307;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
Db 150 FGPKSF 155

RESULT 9

Probable cytochrome P450 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 20-Apr-2001
C:Accession: F86441
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;
ansen, C.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzilli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86441
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-593 <STO>
A:Cross-references: GB:AE005172; NID:911336728; PIDN:AAG31309.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: pea cytochrome P450 CYP97; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:514/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 86.1%; Score 31; DB 2; Length 593;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FGPKNF 6
Db 145 FGPKSF 150

RESULT 10

cysteine-rich intestinal protein - rat
GYRTI
N:Alternate names: CRIP
C:Species: Rattus norvegicus (Norway rat)
C>Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 22-Jun-1999
C:Accession: A03270; A56412
R:Birkemeier, E.H.; Gordon, J.I.
Proc. Natl. Acad. Sci. U.S.A. 83, 2516-2520, 1986
A>Title: Developmental regulation of a gene that encodes a cysteine-rich intestinal p
A:Reference number: A03270; MUID:86205983; PMID:3085096
A:Accession: A03270

A:Molecule type: mRNA
A:Residues: 1-77 <HIR>
A:Cross-references: GB:M13018; NID:9192725; PIDN:AAA37458.1; PID:9309189
A:Experimental source: strain Sprague-Dawley
R:Hemp, J.M.; Cousins, R.J.
Proc. Natl. Acad. Sci. U.S.A. 88, 9671-9674, 1991
A>Title: Cysteine-rich intestinal protein binds zinc during transmembrane zinc transpo
A:Reference number: A56412; MUID:92052153; PMID:1946385
A:Accession: A56412

A:Molecule type: protein
A:Status: preliminary
A:Residues: 4-56 <HEM>
C:Comment: The authors report two short internal duplications, residues 3-27 with 30-
C:Comment: The authors report that the sequence of residues 28-72 is similar to the 1
C:Comment: The concentration in intestinal tissues of the mRNA encoding this protein
C:Comment: Homologous genes for this protein are present in human, bird, fish, and se
C:Superfamily: cysteine-rich intestinal protein; LIM metal-binding repeat homology
C:Keywords: duplication; intestine; intestine; zinc
F:4-55/Domain: LIM metal-binding repeat homology <LIM>

Query Match 83.3%; Score 30; DB 1; Length 77;
Best Local Similarity 83.3%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
Db 61 FGPKGF 66

RESULT 11

JC2431
cysteine-rich protein 1 - human
N:Alternate names: cysteine-rich heart protein; LIM containing protein
C:Species: Homo sapiens (man)
C>Date: 15-Feb-1995 #sequence_revision 26-May-1995 #text_change 20-Aug-1999
C:Accession: JC2431
R:Tsuji, S.K.W.; Yam, N.Y.H.; Lee, C.Y.; Waye, M.M.Y.
Biochem. Biophys. Res. Commun. 205, 497-505, 1994
A>Title: Isolation and characterization of a cDNA that codes for a LIM-containing pro
A:Reference number: JC2431; MUID:95091772; PMID:7999070
A:Accession: JC2431
A:Molecule type: mRNA
A:Residues: 1-77 <TSU>
A:Cross-references: EMBL:U09770; NID:9719268; PIDN:AAA64537.1; PID:9719269
A:Experimental source: heart
C:Genetics:
A:Gene: GDB:CRIP1; CRIP
A:Cross-references: GDB:439373
C:Superfamily: cysteine-rich intestinal protein; LIM metal-binding repeat homology
C:Keywords: cardiac muscle; heart; zinc finger
F:4-55/Domain: LIM metal-binding repeat homology <LIM>

Query Match 83.3%; Score 30; DB 2; Length 77;
 Best Local Similarity 83.3%; Pred. No. 22;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 |||||
 Db 61 FGPKGF 66

RESULT 12
 G02666
 cysteine-rich protein 1 - human
 N:Alternate names: cysteine-rich intestinal protein
 C:Species: Homo sapiens (man)
 C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 20-Aug-1999
 C:Accession: G02666
 R:Blanchard, R.K.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: H01571
 A:Accession: G02666
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-77 <BLA>
 A:Cross-references: EMBL:U58630; NID:g1381637; PIDN:AAB61158.1; PID:g1381638
 C:Genetics:
 A:Gene: GDB:CRIP1; CRIP
 A:Cross-references: GDB:439373
 C:Superfamily: cysteine-rich intestinal protein: LIM metal-binding repeat homology
 F:4-55/Domain: LIM metal-binding repeat homology <LIM2>

Query Match 83.3%; Score 30; DB 2; Length 77;
 Best Local Similarity 83.3%; Pred. No. 22;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 |||||
 Db 61 FGPKGF 66

RESULT 13
 S67982
 hemoglobin beta-III chain - Murana helena
 C:Species: Murana helena
 C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 04-Mar-2000
 C:Accession: S67982
 R:Bellegirini, M.; Giardina, B.; Olinas, A.; Sanna, M.T.; Delana, A.M.; Salvadori, S.; d
 Eur. J. Biochem. 234, 431-436, 1995
 A:Title: Structure/function relationships in the hemoglobin components from moray (Mura
 A:Reference number: S67981; MUID:96128170; PMID:8536685
 A:Accession: S67982
 A:Molecule type: protein
 A:Residues: 1-147 <PEL>
 C:Superfamily: globin: globin homology
 C:Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrie
 F:3-147/Domain: globin homology <GLB>
 F:63/Binding site: oxygen (His) (distal axial ligand) #status predicted
 F:92/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 83.3%; Score 30; DB 2; Length 147;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 |||||
 Db 118 FGPKGF 123

RESULT 14
 S38879
 LIM-domain protein CRP1 - Japanese quail
 C:Species: Coturnix coturnix japonica (Japanese quail)
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 15-Jun-2001
 C:Accession: S38879

R:Welskirchen, R.; Beckerle, M.C.; Bister, K.
 submitted to the EMBL Data Library, November 1993
 A:Description: The CRP family of LIM-domain proteins: identification of two different
 A:Reference number: S38879
 A:Accession: S38879
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-192 <WEI>
 C:Cross-references: EMBL:Z28333; NID:9429159; PID:g1334627
 C:Superfamily: mouse muscle LIM protein MLP: LIM metal-binding repeat homology
 F:10-61/Domain: LIM metal-binding repeat homology <LIM1>
 F:118-169/Domain: LIM metal-binding repeat homology <LIM2>

Query Match 83.3%; Score 30; DB 2; Length 192;
 Best Local Similarity 83.3%; Pred. No. 54;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 |||||
 Db 174 FGPKGF 179

RESULT 15
 A49648
 cysteine-rich protein - chicken
 N:Alternate names: zyxin-binding protein
 C:Species: Gallus gallus (chicken)
 C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 15-Jun-2001
 C:Accession: A49648; S34179
 R:Crawford, A.W.; Pino, J.D.; Beckerle, M.C.
 J. Cell Biol. 124, 117-127, 1994
 A:Title: Biochemical and molecular characterization of the chicken cysteine-rich prot
 A:Reference number: A49648; MUID:94124603; PMID:8294495
 A:Accession: A49648
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-192 <CRA>
 A:Cross-references: EMBL:X73831; NID:g313286; PIDN:CAAS2053.1; PID:g313287
 C:Superfamily: mouse muscle LIM protein MLP: LIM metal-binding repeat homology
 C:Keywords: actin binding; muscle
 F:10-61/Domain: LIM metal-binding repeat homology <LIM1>
 F:118-169/Domain: LIM metal-binding repeat homology <LIM2>

Query Match 83.3%; Score 30; DB 2; Length 192;
 Best Local Similarity 83.3%; Pred. No. 54;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FGPKNF 6
 |||||
 Db 174 FGPKGF 179

Search completed: July 18, 2003, 16:04:26
 Job time : 12.3684 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: July 18, 2003, 15:56:19 : Search time 37.8947 Seconds
(Without alignments)
28.131 Million cell updates/sec

Title: US-10-054-968-8

Perfect score: 35

Sequence: 1 RVAIAGVL 8

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	8	19	AAW4381
2	35	100.0	180	19	AAW4388
3	35	100.0	279	23	ABW47922
4	35	100.0	288	23	ABW47921
5	35	100.0	289	22	AAU34952
6	35	100.0	290	23	ABP38428
7	34	97.1	279	21	AA181709
8	34	97.1	279	22	AAU37847
9	34	97.1	279	22	AAW01016
10	34	97.1	280	23	ABP26997

11	34	97.1	280	23	ABP26998
12	34	97.1	288	23	ABW3584
13	34	97.1	300	21	AAW05954
14	32	91.4	275	22	AAU37848
15	32	91.4	291	22	AAU34951
16	32	91.4	835	22	ABW67504
17	31	88.6	70	31	ABP10464
18	31	88.6	212	22	AAU45973
19	31	88.6	215	22	AAU36069
20	31	88.6	268	23	ABW47338
21	31	88.6	277	31	ABW53583
22	31	88.6	363	22	AAW39366
23	31	88.6	363	22	AAW39366
24	31	88.6	716	22	AAU64250
25	30	85.7	270	23	ABW3216
26	30	85.7	339	20	AAU04961
27	30	85.7	558	23	ABW27849
28	30	85.7	558	23	ABW53631
29	29	82.9	170	21	AAW16455
30	29	82.9	170	21	AAW45317
31	29	82.9	212	22	AAW5505
32	29	82.9	254	21	AAW21904
33	29	82.9	254	21	AAW50938
34	29	82.9	254	21	AAW50938
35	29	82.9	258	21	AAW21903
36	29	82.9	258	21	AAW30708
37	29	82.9	258	21	AAW50937
38	29	82.9	258	21	AAW50974
39	29	82.9	262	21	AAW16454
40	29	82.9	262	21	AAW5316
41	29	82.9	265	21	AAW5315
42	29	82.9	265	21	AAW50936
43	29	82.9	269	22	AAW5782
44	29	82.9	272	21	AAW5227
45	29	82.9	273	21	AAW50973

ALIGNMENTS

RESULT 1
AAW4381
ID AAW4381 standard; peptide: 8 AA.
XX
AC AAW4381;
DT 10-JUN-1998 (first entry)
XX
DE Vancomycin resistant Enterococcus faecium specific epitope SEQ ID NO:8.
XX
KW Vancomycin resistant Enterococcus faecium; VRE; epitope; immunogen;
KW methicillin resistant Staphylococcus aureus; MRSA; transporter protein.
XX
OS Enterococcus faecium.
XX
PN WO9801154-A2.
XX
PD 15-JAN-1998.
XX
PF 07-JUL-1997; 97WO-GB01830.
XX
PR 06-JUL-1996; 96GB-0014274.
XX
FA (NEUT-) NEUTEC PHARM PLC.
XX Burnle JP, Matthews RC;
XX WPI, 1998-100821/09.
XX
XX Treating and diagnosing bacterial and fungal infection with ABC
XX transporter protein - or neutralising or binding agents, and new
XX Staphylococcal proteins, particularly for infections caused by drug
XX resistant Staphylococci and Enterococci

XX Claim 9; Page 39; 62pp; English.
 XX
 CC The present sequence represents a vancomycin resistant Enterococcus
 CC faecium (VRE) specific epitope. The present invention describes the
 CC therapeutic or diagnostic use of a bacterial or fungal ABC transporter
 CC protein, or its immunogenic fragments, in human or veterinary medicine.
 CC Binding agents, specific for an ABC transporter protein, are used in
 CC standard binding assays to detect the ABC transporter protein, while
 CC ABC transporter proteins are used similarly to detect specific
 CC antibodies in patient sera, specifically for diagnosis of Staphylococci
 CC and Enterococci infections. Neutralising agents specific for ABC
 CC transporter proteins, specifically antibodies (Ab), are used to treat
 CC such infections (optionally when coupled to a drug) while ABC
 CC transporter proteins are used as immunogens to protect against
 CC infection. Ab raised against ABC transporter proteins are effective
 CC against multiple drug resistant strains of bacteria (specifically
 CC vancomycin and/or methicillin resistant strains) for which no drug
 CC therapy is available.
 CC
 XX
 SQ Sequence 8 AA;
 Query Match 100.0%; Score 35; DB 19; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 RVAIAGVL 8
 Db 1 RVAIAGVL 8
 RESULT 2
 AAM44388
 ID AAM44388 standard; Protein; 180 AA.
 XX
 AC AAM44388;
 DT 10-JUN-1998 (first entry)
 XX
 DE Vancomycin resistant Enterococcus faecium expression product.
 XX
 KM Vancomycin resistant Enterococcus faecium; VRE; epitope; immunogen;
 KM methicillin resistant Staphylococcus aureus; MRSA; transporter protein.
 XX
 OS Enterococcus faecium.
 XX
 PN WO9801154-A2.
 PD 15-JAN-1998.
 XX
 PF 07-JUL-1997; 97WO-GB01830.
 XX
 PR 06-JUL-1996; 96GB-0014274.
 XX
 PA (NEUT-) NEUTEC PHARM PLC.
 XX
 PI Burnie JP, Mathews RC;
 PI
 DR WPI; 1998-100821/09.
 DR N-PSDB; AAV15410.
 XX
 PT Treating and diagnosing bacterial and fungal infection with ABC
 PT transporter protein - or neutralising or binding agents, and new
 PT Staphylococcal proteins, particularly for infections caused by drug
 PT resistant Staphylococci and Enterococci
 XX
 PS Example; Page 36; 62pp; English.
 XX
 CC The present sequence represents a vancomycin resistant Enterococcus
 CC faecium (VRE) expression product. The present invention describes the
 CC therapeutic or diagnostic use of a bacterial or fungal ABC transporter
 CC protein, or its immunogenic fragments, in human or veterinary medicine.
 CC Binding agents, specific for an ABC transporter protein, are used in

CC standard binding assays to detect the ABC transporter protein, while
 CC ABC transporter proteins are used similarly to detect specific
 CC antibodies in patient sera, specifically for diagnosis of Staphylococci
 CC and Enterococci infections. Neutralising agents specific for ABC
 CC transporter proteins, specifically antibodies (Ab), are used to treat
 CC such infections (optionally when coupled to a drug) while ABC
 CC transporter proteins are used as immunogens to protect against
 CC infection. Ab raised against ABC transporter proteins are effective
 CC against multiple drug resistant strains of bacteria (specifically
 CC vancomycin and/or methicillin resistant strains) for which no drug
 CC therapy is available.
 CC
 XX
 SQ Sequence 180 AA;
 Query Match 100.0%; Score 35; DB 19; Length 180;
 Best Local Similarity 100.0%; Pred. No. 8.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 RVAIAGVL 8
 Db 157 RVAIAGVL 164
 RESULT 3
 ABB47922
 ID ABB47922 standard; Protein; 279 AA.
 XX
 AC ABB47922;
 DT 05-FEB-2002 (first entry)
 XX
 DE Listeria monocytogenes protein #626.
 XX
 KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KM vitamin B12; bacterial infection; disease.
 XX
 OS Listeria monocytogenes.
 XX
 PN WO200177335-A2.
 PD 18-OCT-2001.
 XX
 PF 11-APR-2001; 2001WO-FR01118.
 XX
 PR 11-APR-2000; 2000FR-0004629.
 XX
 PA (INSP) INST PASTEUR.
 XX
 PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fslh H, Dehoux P;
 PI Dussurget O, Chetouani F, Nedjari H, Glaeser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tlierrez-Martinez A, Amend A;
 PI Chakraborty T, Doman E, Hain T, Berche P, Charbit A, Durant L,
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
 PI Madueno E, De Pablos B, Wehlund J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX
 DR WPI; 2002-010914/01.
 DR
 PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides
 XX
 PS Claim 6; SEQ ID No 627; 192pp; French.
 XX
 CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and

CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 279 AA;

Query Match 100.0%; Score 35; DB 23; Length 279;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVAIAGVL 8
 |||||
 Db 148 RVAIAGVL 155

RESULT 4
 ABB47921
 ID ABB47921 standard; Protein; 288 AA.
 XX
 AC ABB47921;
 XX
 DT 05-FEB-2002 (first entry)
 XX
 DE Listeria monocytogenes protein #625.

XX
 KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KM vitamin B12; bacterial infection; disease.
 XX
 OS Listeria monocytogenes.
 OS
 PN WO200177335-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 11-APR-2001; 2001WO-FR01118.
 XX
 PR 11-APR-2000; 2000FR-0004629.

XX
 PA (INSP) INST PASTEUR.

PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fslhi H, Dehoux P;
 PI Dussurget O, Chetouiati F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kieft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Meduendo E, De Pablos B, Wehland J, Kaerst U, Ertian K, Hauf J;
 PI Rose M, Voss H;

WPI; 2002-010914/01.

XX
 DR WPI; 2002-010914/01.
 XX
 PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides

XX
 PS Claim 6; SEQ ID No 626; 192pp; French.

XX
 CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes B2D-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin

CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 288 AA;

Query Match 100.0%; Score 35; DB 23; Length 288;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVAIAGVL 8
 |||||
 Db 153 RVAIAGVL 160

RESULT 5
 AAU34952
 ID AAU34952 standard; Protein; 289 AA.
 XX
 AC AAU34952;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Enterococcus faecalis cellular proliferation protein #239.

XX
 KM Antisense; prokaryotic cellular proliferation protein;
 KM antibiotic; antibacterial; drug design.
 XX
 OS Enterococcus faecalis.
 OS
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.

XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.

XX
 PA (ELITRA) ELITRA PHARM INC.

XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 PI
 XX
 DR WPI; 2001-611495/70.
 XX
 N-PSDB; AAS52811.

XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 PT
 XX
 PS Example 3; Seq ID No 10545; 511pp; English.

XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.

CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 289 AA;

Query Match 100.0%; Score 35; DB 22; Length 289;
Best Local Similarity 100.0%; Pred. No. 14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
Db 153 RVAIAGVL 160

RESULT 6

ABP38428
ID ABP38428 standard; Protein; 290 AA.

AC ABP38428;

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3273.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy.

XX Staphylococcus epidermidis.

XX US6380370-B1.

XX 30-APR-2002.

XX 13-AUG-1998; 98US-0134001.

XX 14-AUG-1997; 97US-055779P.

PR 08-NOV-1997; 97US-064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI, 2002-381255/41.

DR N-PSDB; ABN90973.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX Disclosure; SEQ ID 3273; 267bp; English.

XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.

XX Sequence 290 AA;

Query Match 100.0%; Score 35; DB 23; Length 290;
Best Local Similarity 100.0%; Pred. No. 14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 RVAIAGVL 8
Db 171 RVAIAGVL 178

RESULT 7

AAV81709
ID AAV81709 standard; Protein; 279 AA.

AC AAV81709;

DT 02-JUN-2000 (first entry)

DE Streptococcus pneumoniae protein sequence ID2.

XX Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;
KW bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;
KW kidney disease; diabetes; immunosuppressive disorder; otitis media;
KW pneumococcal septicemia; sinusitis; meningitis; therapy.

XX Streptococcus pneumoniae.

XX WO200006738-A2.

XX 10-FEB-2000.

XX 27-JUL-1999; 99WO-GB02452.

XX 27-JUL-1998; 98GB-0016336.

PR 19-MAR-1999; 99US-0125329.

XX (MICR-) MICROBIAL TECHNICS LTD.

XX Le Page RWF, Wells JM, Hanniffy SB, Hansbro PM;

XX WPI; 2000-195301/17.

DR N-PSDB; AA291805.

XX Streptococcal proteins and polynucleotides useful for diagnosis,
PT treatment and prophylaxis of bacterial infections -
XX Claim 2; Page 40; 76pp; English.

XX This sequence represents a Streptococcus pneumoniae protein of the
CC invention. The proteins (or their homologues, derivatives and/or
CC fragments) are useful as immunogens or antigens. Immunogenic or antigenic
CC compositions comprising the proteins are useful as vaccines and also in
CC diagnostic assays. The sequences are useful for the detection or
CC diagnosis of S. pneumoniae infection, by contacting a sample to be tested
CC with them. Agents capable of antagonising, inhibiting or interfering with
CC the function or expression of the protein or polypeptide are useful in
CC medical compositions in the treatment or prophylaxis of S. pneumoniae
CC infection. As the sequences can be used to treat S. pneumoniae infection,
CC they can be used to treat bacterial pneumonia, which has high rates in
CC young children, the elderly, and in patients with predisposing conditions
CC such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,
CC or with immunosuppressive disorders, especially AIDS. They can also be
CC used to treat pneumococcal septicemia, otitis media, sinusitis, and
CC meningitis.

XX Sequence 279 AA;

Query Match 97.1%; Score 34; DB 21; Length 279;
Best Local Similarity 87.5%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RVAIAGVL 8
Db 153 RVAIAGIL 160

RESULT 8
AAU37847 standard; Protein: 279 AA.
XX
AC AAU37847;
XX
DT 14-FEB-2002 (first entry)
XX
DE Streptococcus pneumoniae cellular proliferation protein #276.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Streptococcus pneumoniae.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr CJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI: 2001-611495/70.
DR N-PSDB; AAS55706.
XX
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 13440; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 279 AA;
XX
XX
XX Query Match 97.1%; Score 34; DB 22; Length 279;
XX Best Local Similarity 87.5%; Pred. No. 22;
XX Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 9
AAM01016 standard; Protein: 279 AA.
XX
AC AAM01016;
XX
DT 02-OCT-2001 (first entry)
XX
DE CFE 15 protein sequence.
XX
KW Antibacterial; vaccine; gene therapy; bacterial cell wall viability;
KW CFE; CEG; Conserved Essential Gene; bacterial infection;
KW antisense therapy; antibiotic resistance.
XX
OS Streptococcus pneumoniae.
XX
PN WO200149721-A2.
XX
PD 12-JUL-2001.
XX
PF 29-DEC-2000; 2000WO-US35604.
XX
PR 30-DEC-1999; 99US-0174089.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Dougherty TJ, Pucci MJ, Dougherty BA, Davison DB, Brucoleri RE;
PI Thanassi JA;
XX
XX WPI: 2001-496721/54.
DR N-PSDB; AAH90715.
XX
XX Nucleic acids encoding conserved essential genes involved in bacterial
PT replication which are potential targets for the treatment of antibiotic
PT resistant bacterial infections -
XX
XX Claim 27; Pages 254-255; 380pp; English.
XX
XX The present invention relates to nucleic acids (AAH90701-AAH90918)
CC encoding polypeptides (AAM01002-AAM01114), which are essential for the
CC viability of a bacterial cell wall. The acronym CFE stands for "Conserved
CC Expression", where CEG stands for "Conserved Essential Gene". The nucleic
CC acids are useful for detecting the presence of proteins essential for the
CC viability of a bacterial cell wall in samples such as cells, tissues,
CC biological fluids, blood, serum, nose, ear or throat swabs with ligands,
CC and for detecting corresponding target nucleic acid molecules with
CC complementary sequences. The nucleic acids are also useful for
CC determining whether a genomic nucleotide sequence of interest is
CC essential for viability of a bacterial cell or whether it resides within
CC an operon, by integrating an exogenous nucleotide sequence comprising a
CC portion of an open reading frame of the genomic sequence of interest
CC (comprising 200-500 base pairs) into the genomic sequence of interest
CC which confers a selectable phenotype to the cell, and determining cell
CC viability with a selection agent such as chloramphenicol. The nucleic
CC acids and proteins are also useful as vaccines and for treating bacterial
CC infections with gene therapy and antisense therapy. The nucleic acids
CC also enable identification of targets suitable for the treatment of
CC antibiotic resistant bacterial infections.
XX
XX Sequence 279 AA;
XX
XX
XX Query Match 97.1%; Score 34; DB 22; Length 279;
XX Best Local Similarity 87.5%; Pred. No. 22;
XX Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 10
ABP26997 standard; Protein: 280 AA.
ID

XX ABP26997;
 AC 02-JUL-2002 (first entry)
 XX
 DT Streptococcus polypeptide SEQ ID NO 3170.
 DE
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus agalactiae.
 OS
 XX W0200234771-A2.
 PN
 XX 02-MAY-2002.
 PD
 XX 29-OCT-2001; 2001WO-GB04789.
 PF
 XX 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 PI
 XX WPI: 2002-352536/38.
 DR N-PSDB; ABN67628.
 DR
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 PS Claim 1; Page 3470; 4525pp; English.
 XX
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 XX
 XX Sequence 280 AA;
 SQ
 OY 1 RVAIAGVL 8
 DB 153 RVAIAGIL 160
 Query Match 97.1%; Score 34; DB 23; Length 280;
 Best Local Similarity 87.5%; Pred. No. 22;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DT 02-JUL-2002 (first entry)
 XX
 XX Streptococcus polypeptide SEQ ID NO 3172.
 DE
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus pyogenes.
 OS
 XX W0200234771-A2.
 PN
 XX 02-MAY-2002.
 PD
 XX 29-OCT-2001; 2001WO-GB04789.
 PF
 XX 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 PI
 XX WPI: 2002-352536/38.
 DR N-PSDB; ABN67629.
 DR
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 PS Claim 1; Page 3470; 4525pp; English.
 XX
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 XX
 XX Sequence 280 AA;
 SQ
 OY 1 RVAIAGVL 8
 DB 153 RVAIAGIL 160
 Query Match 97.1%; Score 34; DB 23; Length 280;
 Best Local Similarity 87.5%; Pred. No. 22;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 11
 ABP26998
 ID ABP26998 standard; Protein; 280 AA.
 XX
 AC ABP26998;
 XX

RESULT 12
 ABB53584
 ID ABB53584 standard; Protein; 288 AA.
 XX
 AC ABB53584;
 XX
 DT 16-MAY-2002 (first entry)
 XX
 DE Lactococcus lactis protein yche.

```

XX  Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
KW  Lactococcus lactis IL1403.
OS  FR2807446-A1.
XX  12-OCT-2001.
XX  11-APR-2000; 2000FR-0004630.
XX  11-APR-2000; 2000FR-0004630.
XX  11-APR-2000; 2000FR-0004630.
XX  (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX  Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX  WPI; 2002-043418/06.
XX  New nucleotide sequence useful in the identification or Lactococcus
XX  lactis and related species -
XX  Claim 6; SEQ ID No 286; 2504bp; French.
XX  The present invention is related to a Lactococcus lactis nucleotide
XX  sequence (ABA90521) and related proteins (ABR53300-ABR55621). The
XX  nucleic acid sequence is useful in the detection and/or amplification of
XX  nucleic acid sequence, particularly to identify Lactococcus lactis or
XX  related species. The proteins of the invention are useful for the
XX  biosynthesis or biodegradation of a composition of interest. The
XX  invention helps research in lactic bacteria, particularly useful in the
XX  production of yogurt and cheese.
XX  Note: The sequence data for this patent is based on equivalent patent
XX  WO200177334 (published 18-OCT-2001) which is available in electronic
XX  format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX  Sequence 288 AA;
SQ
OY  Query Match 97.1%; Score 34; DB 23; Length 288;
    Best Local Similarity 87.5%; Pred. No. 22;
    Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
    1 RVAIAGVL 8
    152 RVAIAGTIL 159
DB
RESULT 13
AAB05954
ID  AAB05954 standard; Protein; 300 AA.
XX  AAB05954;
XX  20-OCT-2000 (first entry)
XX  Protein deduced from Mycoplasma hyopneumoniae genomic clone pAD913.
XX  Mycoplasma hyopneumoniae; vaccine; antigen; antimicrobial;
XX  enzootic pneumonia.
XX  Mycoplasma hyopneumoniae.
XX  OS  Mycoplasma hyopneumoniae.
XX  Key location/Qualifiers
XX  MISC-difference 224 /note= "Encoded by TGA"
XX  WO200031115-A1.
XX  02-JUN-2000.
XX  19-NOV-1999; 99WO-AU01035.
XX  20-NOV-1998; 98AU-0007273.

```

```

XX  (CSTR ) COMMONWEALTH SCI & IND RES ORG.
XX  Moore RJ, Doran TJ;
XX  WPI; 2000-400031/34.
XX  N-PSDB; AAA56840.
XX  Identifying antigenic polypeptides for use in vaccines against
XX  Mycoplasma infection comprises screening proteins co-expressed with a
XX  marker -
XX  Claim 27; Fig 21; 241pp; English.
XX  The present sequence is encoded by a clone selected from a Mycoplasma
XX  hyopneumoniae genomic DNA library. M. hyopneumoniae causes enzootic
XX  pneumonia in pigs. The infection rarely causes death but often results in
XX  significant depression, leading to reduced weight gain. The animals are
XX  also prone to secondary infection by opportunistic pathogens. Genomic
XX  DNA from M. hyopneumoniae was partially digested with restriction
XX  enzymes and fragments were ligated downstream of the polyHis tag region
XX  in the expression vector pCI. The ligation mixture was used to transform
XX  competent E. coli cells and polyHis positive clones were selected using
XX  anti-polyHis antibodies. Expression of the recombinant vector
XX  generates PolyHis fusion proteins which are easily identified and
XX  isolated. Pigs were vaccinated with plasmid DNA from polyHis
XX  positive clones in order to identify clones encoding antigenic proteins
XX  which confer protection against M. hyopneumoniae infection. Such
XX  proteins may subsequently be used in the manufacture of effective
XX  vaccines.
XX  Sequence 300 AA;
SQ
OY  Query Match 97.1%; Score 34; DB 21; Length 300;
    Best Local Similarity 87.5%; Pred. No. 24;
    Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
    1 RVAIAGVL 8
    166 RVAIAGTIL 173
DB
RESULT 14
AAU37848
ID  AAU37848 standard; Protein; 275 AA.
XX  AAU37848;
XX  14-FEB-2002 (first entry)
XX  Streptococcus pneumoniae cellular proliferation protein #277.
XX  Antisense; prokaryotic cellular proliferation protein;
XX  antibiotic; antibacterial; drug design.
XX  Streptococcus pneumoniae.
XX  WO200170955-A2.
XX  27-SEP-2001.
XX  21-MAR-2001; 2001WO-US09180.
XX  21-MAR-2000; 2000US-191078P.
XX  23-MAY-2000; 2000US-206848P.
XX  26-MAY-2000; 2000US-207727P.
XX  23-OCT-2000; 2000US-242578P.
XX  27-NOV-2000; 2000US-253625P.
XX  22-DEC-2000; 2000US-257931P.
XX  16-FEB-2001; 2001US-269308P.
XX  (ELIT-) ELITRA PHARM INC.

```

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 DR WPI: 2001-611495/70.
 DR N-PSDB: AASS5707.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 13441; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 275 AA;
 Query Match 91.4%; Score 32; DB 22; Length 275;
 Best Local Similarity 87.5%; Pred. No. 58;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RVAIAGVL 8
 DB 148 RVAIAGV 155
 |||||||
 148 RVAIAGV 155
 RESULT 15
 AAU34951
 ID AAU34951 standard; Protein; 291 AA.
 XX
 AC AAU34951;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Enterococcus faecalis cellular proliferation protein #238.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX
 OS Enterococcus faecalis.
 XX
 PN WO200170955-A2.
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001MO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 DR WPI: 2001-611495/70.
 DR N-PSDB: AASS2810.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 10544; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 291 AA;
 Query Match 91.4%; Score 32; DB 22; Length 291;
 Best Local Similarity 87.5%; Pred. No. 62;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RVAIAGVL 8
 DB 160 RVAIAGV 167
 |||||||
 160 RVAIAGV 167

Search completed: July 18, 2003, 16:01:52
 Job time : 38.8947 secs